

Master 1 – Économétrie et Statistiques, parcours Économétrie
Appliquée

Techniques de prévision et conjoncture - Série non saisonnière mensuelle

**Analyse de la série mensuelle de l'indice CAC 40 sur la
période 2000-2018 et prévision pour l'année 2019**

Dossier réalisé par :

QUINTIN DE KERCADIO Pierre
CROCHET Florian

Année universitaire 2024–2025

Résumé

Dans un contexte de forte volatilité des marchés financiers, ce dossier vise à anticiper l'évolution de l'indice boursier CAC 40 sur l'année 2019 à partir de données mensuelles couvrant la période 2000-2018. Nous avons cherché à identifier la méthode de prévision la plus efficace pour modéliser à court terme les fluctuations du CAC 40. Pour y répondre, nous avons mené une analyse rigoureuse en plusieurs étapes : détection de points atypiques, tests de stationnarité, choix d'une transformation logarithmique, puis estimation de plusieurs modèles linéaires. Cinq approches ont été comparées (ARIMA, Holt-Winters, ADAM ETS, ADAM ETS+ARIMA, SSARIMA), évaluées selon les critères AIC, AICc, MSE, R^2_{OOS} , ainsi que le test de Diebold-Mariano. Les résultats montrent que le modèle ADAM ETS+ARIMA offre la meilleure précision prédictive hors échantillon, avec des erreurs quadratiques inférieures à celles du benchmark naïf et des modèles classiques. Ce modèle s'avère robuste, bien adapté aux séries financières non stationnaires, et particulièrement pertinent pour la prévision glissante à un mois.

Mots clés : marchés financiers, prévision à court terme, séries temporelles, modèles linéaires dynamiques, R.

Abstract

In a context of high volatility in financial markets, this report aims to forecast the evolution of the CAC 40 stock index for the year 2019, based on monthly data covering the period from 2000 to 2018. The objective was to identify the most effective forecasting method to model short-term fluctuations in the CAC 40. To this end, we conducted a rigorous multi-step analysis : detection of outliers, stationarity tests, application of a logarithmic transformation, and estimation of several linear models. Five approaches were compared (ARIMA, Holt-Winters, ADAM.ETS, ADAM.ETS+ARIMA, SSARIMA), and assessed using AIC, AICc, MSE, R^2_{OOS} , and the Diebold-Mariano test. The results show that the ADAM.ETS+ARIMA model delivers the best out-of-sample predictive accuracy, with lower squared errors than both the naïve benchmark and classical models. This model proves to be robust, well suited to non-stationary financial series, and particularly effective for one-month rolling forecasts.

Keywords : financial markets, short-term forecasting, time series analysis, dynamic linear models, R.

Sommaire

Table des matières

1	Analyse préliminaire	1
1.1	Présentation de la série	1
1.2	Détection des points atypiques	3
1.3	Vérification de l'absence de saisonnalité et choix du modèle log-niveau . . .	5
1.4	Vérification de la stationnarité de la série	6
1.5	Statistiques descriptives	8
2	Estimation des modèles linéaires	13
2.1	Estimation des paramètres du modèle ARIMA(p,d,q) et de la méthode LED Holt-Winters, ADAM ETS, ADAM ETS ARIMA et SSARIMA . . .	14
2.1.1	ARIMA(1,1,1)	14
2.1.2	Modèle Holt-Winters LED	14
2.1.3	ADAM ETS	16
2.1.4	ADAM ETS + ARIMA	17
2.1.5	SSARIMA(0,1,3)(0,0,3)[12]	18
2.2	Paramètres des modèles précédents et détermination du meilleur modèle d'après les critères AIC et AICc	20
3	Prévision linéaire sur une année avec représentation graphique	22
4	Qualité de prévision	25
4.1	MSE et R^2_{OOS} des modèles précédents	25
4.2	Comparaison avec une prévision naïve	27
4.3	Graphique des CSPE	29

5	Test de précision	32
5.1	Test de Diebold-Mariano	32
5.2	Définition du meilleur modèle	33
6	Prévision sur une année avec un pas de 1 mois	35
7	Conclusion	37
8	Bibliographie	38
9	Annexes	40

1 Analyse préliminaire

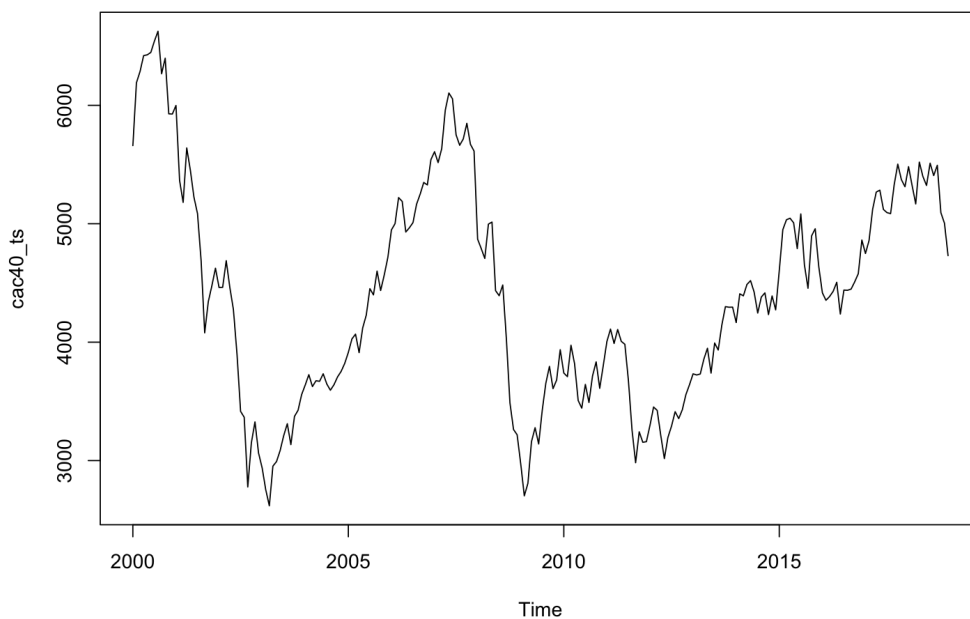
1.1 Présentation de la série

La série de données étudiée correspond à l'évolution de l'indice CAC 40, principal indice boursier de la place de Paris. Celui-ci reflète la performance des 40 plus grandes capitalisations cotées sur le marché Euronext de Paris, pondérées par leur capitalisation flottante. L'indice est calculé et publié en continu par Euronext, et constitue un indicateur clé de la conjoncture économique française et européenne.

Les données ont été récupérées sur la plateforme Yahoo Finance¹, qui fournit des séries quotidiennes, hebdomadaires et mensuelles pour les principaux indices mondiaux. Dans le cadre de cette étude, nous utilisons une fréquence mensuelle, en considérant les valeurs de clôture de l'indice CAC 40 sur la période allant de janvier 2000 à décembre 2018, couvrant ainsi une période de 18 ans.

L'analyse graphique de cette série permettra d'identifier les principales dynamiques structurelles, notamment les phases de croissance, les retournements de tendance, ainsi que les éventuelles ruptures liées à des chocs économiques ou financiers.

FIGURE 1 – Évolution de la série du CAC 40 entre 2000 et 2018



1. Yahoo Finance. (2025). *Finance news, data, and tools*

Le graphique ci-dessus présente l'évolution du CAC 40 sur la période étudiée. Il met en évidence plusieurs phases économiques caractéristiques, notamment les périodes de crise, de reprise ou de stagnation. Nous pouvons remarquer une période de baisse entre 2000 et 2003 due à l'éclatement de la bulle internet. L'explosion de cette bulle spéculative liée aux entreprises technologiques, les *dot-com bubble*, a provoqué un effondrement global des marchés².

Par la suite, durant la période 2003 à 2007, le CAC 40 connaît une forte reprise allant jusqu'à 6 000 points, portée par un contexte de croissance mondiale soutenue, notamment dans les pays émergents (BRICS), et par des politiques monétaires favorables³.

Cependant, entre 2007 et 2009, le CAC 40 s'effondre sous l'effet de la crise des subprimes, déclenchée par des défauts massifs sur les crédits immobiliers américains et amplifiée par la faillite de Lehman Brothers⁴.

Dans un second temps, entre 2009 et 2012, l'indice connaît une reprise progressive grâce aux plans de relance et aux politiques monétaires expansionnistes, avant d'être freiné par la crise des dettes souveraines en zone euro⁵.

Puis, entre les années 2012 et 2015, le CAC 40 remonte progressivement vers 5 000 points, soutenu par l'intervention déterminante de la Banque centrale européenne. En affirmant qu'il ferait "whatever it takes" pour sauver l'euro, Mario Draghi a restauré la confiance des marchés, appuyée par une politique monétaire extrêmement accommodante⁶.

Ensuite, de 2015 à 2018, le CAC 40 progresse de manière irrégulière dans un contexte de reprise économique mondiale synchronisée, de faibles taux d'intérêt et d'amélioration des résultats des entreprises. En 2017, les marchés financiers affichent un regain d'optimisme et un retour de l'appétence pour le risque, stimulé notamment par l'élection du Président Emmanuel Macron⁷.

Enfin, en 2018, le CAC 40 enregistre une baisse de 10,95 %, sa plus forte chute annuelle depuis 2008, en raison des tensions commerciales, du ralentissement économique mondial

2. De Jonghe, A. (2017). *La bulle Internet : origines, mécanismes, conséquences*.

3. Capital. (2013). *Panorama de la Bourse de Paris, 10 ans après la fin de la bulle Internet*.

4. Le Figaro. (2009). *CAC 40, d'un 9 mars à un autre...*

5. Les Échos. (2010). *La zone euro secouée par une grave crise de la dette*.

6. Les Échos. (2013). *Mario Draghi démine le terrain*.

7. Le Monde. (2017). *En 2017, les marchés financiers ont retrouvé le goût du risque*.

et de l'instabilité politique en France⁸.

Il convient dès lors d'identifier les points atypiques de la série, tels que les outliers ou ruptures structurelles, afin de modéliser de manière rigoureuse les dynamiques temporelles de l'évolution du CAC 40.

1.2 Détection des points atypiques

La présence de points atypiques, qu'ils soient de nature transitoire ou permanente, constitue une difficulté majeure dans l'analyse des séries temporelles. Ces anomalies peuvent fortement perturber l'estimation des modèles en introduisant un biais. Pour résoudre ce problème, nous nous appuyerons sur l'approche de Chen et Liu (1993). Cette méthode identifie automatiquement les chocs ponctuels ou structurels dans la série (AO, LS, TC, IO, SO), en estimant un modèle ARIMA avec interventions, puis fournit une série ajustée, corrigée des effets de ces anomalies.

Il existe cinq types de points atypiques susceptibles d'influencer une série temporelle. Dans le cas présent, plusieurs points atypiques ont été identifiés dans la série du CAC40, parmi lesquels un Additive Outlier (AO), un Level Shift (LS) et un Temporary Change (TC).

Tout d'abord, un *Additive Outlier* (AO) a été détecté. Ce type de point atypique affecte une seule observation à un moment précis dans la série temporelle, entraînant ainsi une perturbation localisée et isolée. Ce phénomène se traduit par un décalage instantané, sans impact durable sur la dynamique générale de la série :

$$\xi_{AO} = 1$$

En revanche, un *Level Shift* (LS) a été observé, impliquant un changement permanent du niveau de la série. Contrairement à l'*Additive Outlier*, cette perturbation ne concerne pas une seule observation, mais affecte de manière durable toutes les périodes suivantes, modifiant le niveau global de la série de façon continue. Cette rupture peut être modélisée

8. Le Parisien. (2019). *Bourse : le CAC 40 termine 2018 en baisse de près de 11 %*.

par l'expression suivante :

$$\xi_{LS} = \frac{1}{1 - L}, \quad \text{où } L \text{ représente l'opérateur de retard.}$$

Enfin, un *Temporary Change* (TC) a été identifié. Ce type de point atypique génère un effet temporaire, perturbant la série avant qu'elle ne revienne progressivement à son niveau initial. La vitesse à laquelle la série revient à son niveau précédent est contrôlée par le paramètre δ , qui détermine la rapidité du retour exponentiel de la série vers son état initial. Ce phénomène est modélisé par l'expression suivante :

$$\xi_{TC} = \frac{1}{1 - \delta L}, \quad \text{avec } 0 < \delta < 1.$$

Ces points atypiques ont un impact significatif sur la dynamique de la série. Ainsi, le tableau suivant présente les points atypiques identifiés dans la série du CAC40, en précisant leur rang d'observation, leur date d'occurrence, leur nature ainsi que les statistiques associées (t-statistique et p-value).

TABLE 1 – Points atypiques détectés dans la série CAC40

Observation	Date	Type de point	t-stat	p-value
21	09/2001	Temporary Change (TC)	-3,197	0,00158
33	09/2002	Additive Outlier (AO)	-5,214	$4,16 \times 10^{-7}$
97	01/2008	Level Shift (LS)	-3,091	0,00224

Dans la Table 1 ci-dessus, trois points atypiques ont été détectés au sein de la série CAC 40 (voir annexe 1).

Tout d'abord, en septembre 2001, le modèle met en évidence un changement transitoire de niveau (TC) avec une p-value de 0,00158, indiquant un effet significatif au seuil de risque de 1 %. Cette perturbation temporaire, qui s'estompe avec le temps, coïncide avec les attentats du 11 septembre 2001, lesquels ont entraîné une chute brutale mais temporaire des marchés financiers mondiaux. L'incertitude et la panique générées par ce choc géopolitique ont provoqué une réaction immédiate des investisseurs, rapidement

corrigée dans les mois suivants⁹.

Ensuite, en septembre 2002, un point atypique (AO) est détecté, avec une p-value de $4,16 \times 10^{-7}$. Ce point atypique significatif au seuil de risque de 0,1 % signale un choc ponctuel affectant directement la série, sans effet persistant. Cette irrégularité pourrait être liée à une combinaison d'effets postérieurs au 11 septembre, amplifiés par la révélation de scandales financiers majeurs, tels qu'Enron (fin 2001)¹⁰ et WorldCom (mi-2002)¹¹. Ces événements ont effectivement altéré la confiance des investisseurs et contribué à la volatilité accrue des marchés.

Enfin, en janvier 2008, un changement de niveau (LS) est identifié (avec une p-value de 0,00224), traduisant un effet durable sur la série. Ce changement structurel, significatif au seuil de risque de 1 %, reflète la diffusion mondiale de la crise des subprimes. Bien que les premiers signes de tensions sur les marchés immobiliers américains soient apparus dès 2007, c'est au tournant de 2008 que la crise a pris une ampleur systémique, affectant l'ensemble du secteur bancaire mondial. Cette rupture de tendance dans le CAC 40 marque ainsi le début d'une période prolongée de turbulences financières¹².

1.3 Vérification de l'absence de saisonnalité et choix du modèle log-niveau

TABLE 2 – Résultats des tests de saisonnalité et du test log-niveau

Test de saisonnalité				Test log-niveau
Webel-Ollech		Seasonal dummies		
stat	p-value	stat	p-value	
0	0,113	1,55	0,116	logarithme

Les résultats des tests statistiques indiquent une absence de saisonnalité significative dans la série du CAC 40. Le test de Webel-Ollech, dont l'hypothèse nulle postule l'absence de saisonnalité, donne une statistique égale à zéro, accompagnée d'une p-value de 0,113,

9. Les Echos. (2001, septembre). *Le CAC 40 terrassé par les attentats*.

10. Wikipédia. *Scandale Enron*.

11. Wikipédia. *Scandale financier WorldCom de 2002*.

12. Le Figaro. (21 janvier 2008). *La Bourse de Paris dévisse sous les points*.

ce qui ne permet pas de rejeter cette hypothèse au seuil de 5 %. De même, le test basé sur les variables fictives saisonnières (seasonal dummies), qui teste la nullité conjointe des coefficients associés aux modalités saisonnières, fournit une statistique de 1,55 avec une p-value de 0,116, indiquant également la non-significativité au seuil de 5 %. Ces résultats convergents permettent de conclure à l'absence de composante saisonnière significative dans la série temporelle du CAC 40.

Le modèle RegARIMA, mis en œuvre dans la procédure X13, combine une composante de régression et une composante ARIMA. Il a été appliqué à la série du CAC 40 après une transformation logarithmique (annexe 2). Cette transformation est généralement recommandée lorsque les données présentent une variance non constante (hétéroscédasticité), c'est-à-dire lorsque la dispersion des valeurs augmente avec le niveau des observations. Dans le cas d'un indice boursier comme le CAC 40, susceptible d'évoluer sur plusieurs ordres de grandeur, elle permet de stabiliser la variance et de rapprocher la série de la stationnarité, en rendant les moments (moyenne et variance) plus stables au fil du temps. Elle facilite également l'interprétation des coefficients en transformant les effets multiplicatifs en effets additifs. L'utilisation de cette transformation est donc justifiée, car elle contribue à améliorer la qualité des ajustements du modèle.

1.4 Vérification de la stationnarité de la série

Une série temporelle $\{y_t\}$ est dite stationnaire si ses propriétés statistiques ne varient pas dans le temps. On distingue deux notions de stationnarité. La stationnarité au sens faible exige que l'espérance soit constante $E[y_t] = \mu \forall t$, la variance finie et constante $Var(y_t) = \sigma^2 < \infty \forall t$, et que la covariance entre deux observations ne dépende que du décalage s : $Cov(y_t, y_{t-s}) = \gamma_s \forall t, \forall s$. La stationnarité au sens fort est plus exigeante : elle impose que la loi jointe de $\{y_t, y_{t+1}, \dots, y_{t+h}\}$ soit identique à celle de $\{y_s, y_{s+1}, \dots, y_{s+h}\}$ pour tout t, s , c'est-à-dire qu'elle ne dépende que du décalage h . Autrement dit, la distribution jointe de toute sous-séquence ne doit pas dépendre du temps.

La stationnarité est une propriété essentielle en économétrie des séries temporelles, car elle garantit la stabilité des relations statistiques au cours du temps. Cette stabilité permet une estimation fiable et une interprétation cohérente des modèles, notamment

les processus ARMA. Lorsqu'une série est stationnaire, ses caractéristiques — moyenne, variance et structure d'autocorrélation — restent constantes, ce qui renforce la validité des prévisions. En revanche, une série non stationnaire, marquée par des tendances, des ruptures ou des effets saisonniers, peut fausser l'analyse. Il est alors souvent nécessaire de la transformer, par exemple par différenciation, afin de la rendre stationnaire avant toute modélisation.

Pour évaluer la stationnarité de la série, nous appliquons trois tests complémentaires : le test de Dickey-Fuller augmenté (ADF), le test de Phillips-Perron (PP) et le test KPSS. Les tests ADF et PP ont pour hypothèse nulle la présence d'une racine unitaire, c'est-à-dire que la série est non stationnaire. Si cette hypothèse est rejetée, cela suggère que la série est stationnaire. Le test de Phillips-Perron repose sur une méthodologie différente de celle du test ADF, notamment dans la correction de l'autocorrélation et de l'hétéroscédasticité des résidus. En complément, le test KPSS teste l'hypothèse nulle inverse : il considère que la série est stationnaire, et cherche à détecter une non-stationnarité. L'utilisation conjointe de ces trois tests permet ainsi une évaluation robuste de la stationnarité de la série.

TABLE 3 – Résultats des tests de stationnarité avant et après différenciation

Test	Critère	Avant différenciation	Après différenciation
ADF	Statistique de test	-2,1351	-10,759
	Valeur critique 5%	-2,88	-1,95
	Résultat	Non stationnaire	Stationnaire
PP	Statistique de test	-2,1828	-13,1645
	Valeur critique 5%	-2,874	-2,874
	Résultat	Non stationnaire	Stationnaire
KPSS	Statistique de test	0,3004	0,1174
	Valeur critique 5%	0,463	0,463
	Résultat	Stationnaire	Stationnaire

Les résultats des trois tests sont contradictoires avant différenciation. Les tests ADF et PP présentent des statistiques respectives de -2,1351 et -2,1828, supérieures à leurs valeurs critiques à 5 % (-2,88 et -2,874), ce qui conduit à ne pas rejeter l'hypothèse nulle de non-stationnarité : la série semble donc comporter une racine unitaire. En revanche, le

test KPSS fournit une statistique de 0,3004, inférieure à sa valeur critique de 0.463, ce qui ne permet pas de rejeter l'hypothèse nulle de stationnarité. Cette divergence suggère une incertitude sur la nature stationnaire de la série, possiblement liée à une proximité d'un processus de racine unitaire ou à un bruit élevé affectant la qualité des tests.

Après différenciation, les résultats deviennent convergents : les tests ADF et PP rejettent l'hypothèse nulle de non-stationnarité, avec des statistiques de $-10,759$ et $-13,1645$, nettement inférieures aux valeurs critiques à 5 %. Le test KPSS, quant à lui, donne une statistique de 0,1174, toujours inférieure à la valeur critique, et ne permet pas de rejeter l'hypothèse nulle de stationnarité. Ainsi, l'ensemble des tests indique que la série est stationnaire après différenciation.

On en conclut que la série est intégrée d'ordre 1, notée $I(1)$. Autrement dit, elle devient stationnaire après une différenciation d'ordre 1, ce qui est une condition préalable essentielle pour procéder à des modélisations économétriques dynamiques.

1.5 Statistiques descriptives

Après avoir évalué la stationnarité de la série logarithmique à l'aide des tests ADF, PP et KPSS, il convient désormais d'en examiner les propriétés statistiques. Le tableau suivant présente les principales statistiques descriptives de la série logarithmique du CAC 40.

TABLE 4 – Statistiques descriptives de la série logarithmique du CAC 40

Statistique	Valeur
Nombre d'observations	228
Valeurs manquantes (NAs)	0
Minimum	7,739
Maximum	8,659
1 ^{er} quartile	8,146
3 ^e quartile	8,489
Moyenne	8,310
Médiane	8,339
Somme	1894,708
Erreur standard de la moyenne (SE Mean)	0,0137
Intervalle de confiance (95%) de la moyenne	[8,283 ; 8,337]
Variance	0,0428
Écart-type (Stdev)	0,2069
Asymétrie (Skewness)	-0,412
Aplatissement (Kurtosis)	-0,667

La série logarithmique du CAC 40 comprend 228 observations sans valeur manquante. Les statistiques descriptives révèlent une distribution globalement centrée, avec une moyenne de 8,31 et une médiane de 8,34. La moyenne étant légèrement inférieure à la médiane, cela suggère une faible asymétrie négative, traduisant une légère tendance vers des valeurs plus faibles, sans effet notable sur la tendance centrale.

L'écart-type de 0,21, associé à une amplitude allant de 7,74 à 8,66, reflète une dispersion modérée des niveaux logarithmiques du CAC 40 sur la période étudiée. Cette variabilité est cohérente avec les fluctuations structurelles des indices boursiers, influencées par les cycles économiques et financiers.

L'asymétrie (skewness) de -0,41 confirme une distribution légèrement inclinée vers la gauche, indiquant que les valeurs plus faibles ont été un peu plus fréquentes. Toutefois, cette asymétrie reste modérée et ne remet pas en cause l'équilibre global de la distribution.

L'aplatissement (kurtosis) de -0,67 indique une distribution légèrement moins 'pointue'

qu'une distribution normale. Cela implique que les observations sont davantage concentrées autour de la moyenne, avec des queues moins épaisses. Ce profil de distribution est courant pour des séries transformées logarithmiquement, qui atténuent l'influence des valeurs extrêmes.

L'intervalle de confiance à 95 % pour la moyenne, estimé à $[8,28 ; 8,34]$, indique une estimation précise de cette dernière, avec une faible incertitude. Toutefois, cet intervalle ne permet pas de statuer sur la présence ou non d'une tendance dans la série, les log-niveaux pouvant intégrer une tendance stochastique. Seuls les tests ADF, PP ou KPSS effectués précédemment permettent d'en juger de manière rigoureuse.

Afin de compléter l'analyse descriptive de la série logarithmique du CAC 40, un test de normalité de Shapiro-Wilk a été réalisé. Ce test permet d'évaluer dans quelle mesure la distribution observée s'écarte d'une distribution normale théorique. Compte tenu des légères déviations observées au niveau de l'asymétrie et de l'aplatissement, il est pertinent de vérifier statistiquement la normalité de la série. Le tableau ci-dessous présente les résultats obtenus.

TABLE 5 – Résultat du test de normalité de Shapiro-Wilk

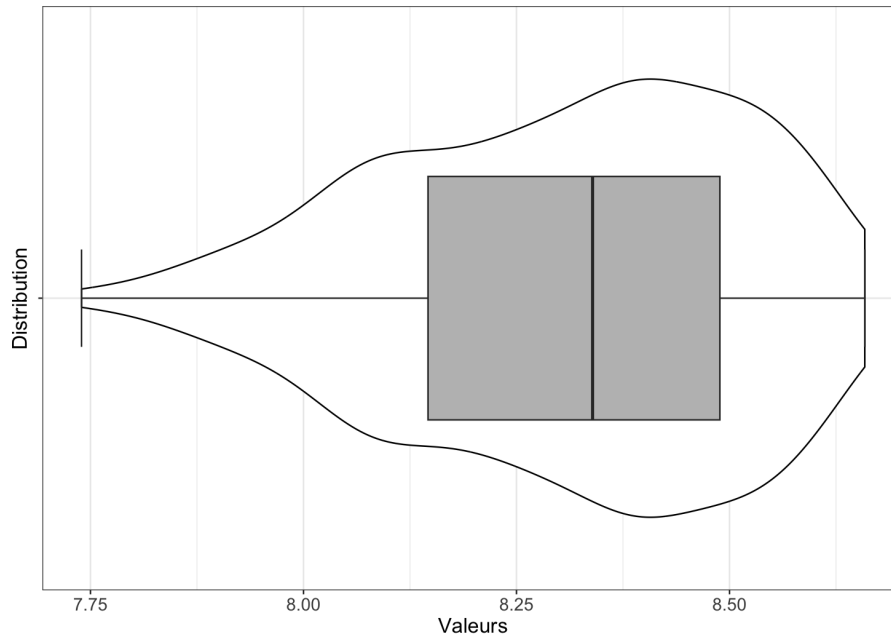
Test de normalité de Shapiro-Wilk	
W	p-value
0,96739	$4,259 \times 10^{-5}$

Les résultats du test de Shapiro-Wilk indiquent une statistique $W = 0,96739$ et une p-value très faible de $4,259 \times 10^{-5}$, largement inférieure au seuil de signification de 1 %. Cette p-value conduit au rejet de l'hypothèse nulle selon laquelle les données suivent une distribution normale. Ainsi, bien que les statistiques descriptives suggèrent une distribution relativement symétrique et modérément aplatie, la normalité est statistiquement rejetée au seuil de 1 % selon ce test.

Pour compléter l'analyse numérique, nous présentons ci-dessous la boîte à moustaches et la densité de distribution de la série logarithmique du CAC 40. Ce graphique synthétise visuellement la dispersion, l'asymétrie et l'aplatissement de la distribution, tout en illus-

trant les écarts à la normalité mis en évidence par le test de Shapiro-Wilk. Il constitue ainsi un complément graphique aux propriétés statistiques précédemment décrites.

FIGURE 2 – Visualisation de la boîte à moustaches et de la distribution de la série logarithmique du CAC 40



Le graphique ci-dessus illustre la boîte à moustaches associée à la distribution de la série logarithmique du CAC 40. Ce type de représentation, combinant une boîte à moustaches à une densité (sous forme de violon), permet d’apprécier simultanément la dispersion, la symétrie et la concentration des valeurs. Il donne une vue d’ensemble de la structure de la distribution, tout en confirmant l’absence de valeurs atypiques dans cette série temporelle préalablement nettoyée.

Sur le plan de la distribution statistique, la boîte centrale représente l’intervalle interquartile (IQR), allant du premier au troisième quartile (Q1 à Q3), et contient les 50 % des observations centrales. La médiane, indiquée par une ligne noire au centre de la boîte, se situe à 8,34. La moyenne, bien qu’absente du graphique, est estimée à 8,31. Ce léger écart entre moyenne et médiane traduit une asymétrie négative modérée, confirmée par une valeur de skewness de -0,41. Visuellement, la boîte apparaît légèrement décalée vers la droite, mais cette impression est contrebalancée par une extension plus marquée de la queue gauche de la distribution. Cette dissymétrie indique que quelques valeurs faibles, bien que rares, influencent la forme générale de la distribution.

La forme du violon (la courbe de densité) renforce cette lecture : elle montre une forte concentration des observations autour de la médiane, tandis que la densité décroît plus lentement vers les faibles valeurs. Cette queue gauche plus étirée reflète la présence occasionnelle de baisses plus importantes — un comportement caractéristique des séries financières, souvent sensibles à des chocs exogènes.

Les moustaches s'étendent jusqu'à 1,5 fois l'IQR de part et d'autre de la boîte. L'absence de points au-delà de ces limites indique qu'aucune valeur n'est identifiée comme aberrante dans la série logarithmique analysée.

En résumé, ce graphique confirme les statistiques descriptives obtenues : la distribution est peu dispersée, modérément aplatie (kurtosis de $-0,67$) et présente une légère asymétrie à gauche (skewness de $-0,41$), perceptible par une queue plus étendue du côté des faibles valeurs. Ces caractéristiques décrivent une distribution structurée, mais s'écartant des critères classiques de normalité, notamment en raison de sa dissymétrie et de sa faible concentration autour de la moyenne. Cette lecture visuelle est cohérente avec les résultats du test de Shapiro-Wilk, qui rejette significativement l'hypothèse de normalité au seuil de 1 % ($p < 0,01$), confirmant que la distribution logarithmique du CAC 40 ne suit pas une loi normale.

2 Estimation des modèles linéaires

Dans cette section, nous estimons et analysons les paramètres des principaux modèles de prévision applicables à la série étudiée. Il s’agit notamment du modèle classique $\text{ARIMA}(p, d, q)$, de la méthode de lissage exponentiel de Holt-Winters, ainsi que d’approches plus récentes issues du package `smooth`, telles que les modèles ADAM ETS, ADAM ETS avec composante ARIMA, et SSARIMA.

Ayant choisi de travailler directement sur la série logarithmique du cours du CAC 40 — laquelle conserve une dynamique non stationnaire typique des séries de prix financiers — et non sur ses rendements, stationnaires grâce à la différenciation de la série en niveau, nous avons écarté les modèles autorégressifs simples de type $\text{AR}(1)$ ou $\text{AR}(p)$, ainsi que la méthode CES (Complex Exponential Smoothing). Ces modèles reposent sur l’hypothèse de stationnarité de la série, ce qui n’est pas le cas ici : la transformation logarithmique atténue certaines hétéroscédasticités, mais ne supprime pas la tendance stochastique sous-jacente. Le modèle CES, en particulier, bien qu’il permette de modéliser des cycles complexes grâce à une formulation trigonométrique, suppose une stationnarité structurelle et s’avère inadapté aux séries présentant une tendance aléatoire persistante, comme c’est le cas pour le cours du CAC 40, même exprimé en logarithme. Ces approches conviennent ainsi davantage à l’estimation ou à la prévision des rendements, stationnalisables, qu’à celle du cours de l’indice. Or, l’objectif de cette étude est de prévoir directement le niveau logarithmique de l’indice CAC 40, et non ses variations relatives. Dans ce cadre, les modèles retenus — tels que $\text{ARIMA}(p, d, q)$, Holt-Winters (LED), ADAM ETS, ADAM ETS ARIMA ou SSARIMA — intègrent des mécanismes adaptés au traitement de la non-stationnarité (notamment via la différenciation, l’utilisation de composantes d’erreurs ou des structures d’états), et s’avèrent donc pertinents pour notre démarche d’estimation puis de prévision.

Pour chacun de ces modèles, les paramètres estimés seront interprétés, puis comparés à l’aide des critères d’information AIC et AICc, en vue d’identifier la spécification la plus adaptée à la dynamique du logarithme du CAC 40. Cette analyse comparative constituera un point de départ pour la phase de prévision développée dans la deuxième partie.

2.1 Estimation des paramètres du modèle ARIMA(p,d,q) et de la méthode LED Holt-Winters, ADAM ETS, ADAM ETS ARIMA et SSARIMA

2.1.1 ARIMA(1,1,1)

Le modèle ARIMA(1,1,1) a été sélectionné automatiquement sur la série logarithmique `log_cac40_ts` à l'aide de la fonction `auto.arima`. Ce modèle comprend une différenciation d'ordre 1, ce qui signifie que la série est stationnaire après différenciation. Il inclut également une composante de moyenne mobile et un terme autorégressif.

Les coefficients estimés révèlent une structure mixte avec un terme autorégressif significatif ($AR(1) = -0,4494$) et un terme de moyenne mobile significatif ($MA(1) = 0,6325$), traduisant une réponse ajustée aux chocs passés. Le terme constant est estimé à $\mu = 0$, ce qui suggère que la moyenne de la série est stable après différenciation.

Le critère AIC associé à ce modèle est de $-756,35$, et l'AICc de $-756,24$, ce qui atteste d'une qualité d'ajustement correcte, bien que supérieure à celle du modèle Holt-Winters. L'erreur quadratique moyenne (RMSE) est de $0,0450$ et l'erreur absolue moyenne (MAE) de $0,0351$, ce qui confirme une capacité de prévision satisfaisante dans un cadre non saisonnier (voir annexe 5).

2.1.2 Modèle Holt-Winters LED

Les méthodes de Holt-Winters et le lissage exponentiel double (LED) sont très proches, et peuvent même être rendues équivalentes en choisissant des paramètres de lissage particuliers. Plus précisément, les deux approches donnent les mêmes résultats si l'on impose les relations suivantes :

$$\alpha = \beta^2, \quad \gamma = 1 - \left(\frac{1 - \beta}{1 - \beta^2} \right)^2 = \frac{2\beta}{1 - \beta^2} \quad (1)$$

Cela permet d'interpréter le modèle de Holt-Winters (dans sa version sans saisonnalité) comme une généralisation du LED, dans laquelle le niveau et la tendance sont ajustés séparément grâce à deux paramètres de lissage.

Cette méthode offre plus de souplesse pour adapter le modèle aux données. On peut, par exemple, rendre le modèle très réactif aux nouvelles observations en choisissant des valeurs de α et γ proches de 1. Dans ce cas, les prévisions suivent de très près les derniers points observés : on parle alors de prévision lisse. À l'inverse, si β et γ sont plus petits, cela revient à donner plus de poids au passé plus ancien, on obtient alors un lissage plus fort.

Le choix des coefficients de lissage peut se faire manuellement, mais dans la pratique, on les estime souvent automatiquement à partir des données, en minimisant l'erreur entre les valeurs prédites et les observations.

Enfin, la prévision à un horizon h se fait simplement en prolongeant la droite ajustée à la dernière période observée. Elle est donnée par la formule suivante :

$$\hat{X}_T(h) = \hat{a}_T + \hat{b}_T h \quad (2)$$

où \hat{a}_T est le niveau estimé à la date T , et \hat{b}_T la pente. Cette équation traduit l'idée que le modèle LED prévoit une évolution linéaire à partir du dernier point connu.

À partir de ce cadre théorique, nous avons estimé le modèle Holt-Winters LED sur la série logarithmique du CAC 40. Ce modèle utilise deux paramètres de lissage indépendants pour mettre à jour le niveau et la tendance, ce qui permet de s'adapter aux évolutions de la série.

TABLE 6 – Paramètres du modèle Holt-Winters LED : α , β , niveau initial et tendance initiale

Paramètre	Valeur estimée
α	1.000
β	0.1795
\hat{a}_0 (niveau initial)	8.462
\hat{b}_0 (tendance initiale)	-0.019

L'estimation automatique réalisée via la fonction `HoltWinters()` fournit les coefficients de lissage suivants : $\alpha = 1$ et $\beta = 0,1795$. Une valeur d' α proche de 1 indique que le niveau est fortement influencé par les observations les plus récentes. La valeur plus modé-

rée de β traduit une mise à jour plus progressive de la pente. Le niveau initial est estimé à $\hat{a}_0 = 8,462$ et la pente initiale à $\hat{b}_0 = -0,019$, ce qui suggère une tendance légèrement décroissante à moyen terme.

La prévision sur un horizon de 12 mois produite par ce modèle affiche une trajectoire globalement décroissante, avec une valeur prévue de 8,443 en décembre 2019 et de 8,242 en novembre 2020. Cette dynamique est cohérente avec la pente négative estimée. Comme attendu avec les méthodes de lissage non amorties, les intervalles de confiance s'élargissent au fur et à mesure de l'horizon, traduisant une incertitude croissante des prédictions. Cela reflète l'effet cumulatif des erreurs de prévision, caractéristique des modèles linéaires non stationnaires en tendance (voir annexe 6).

2.1.3 ADAM ETS

Le modèle ADAM ETS sélectionné automatiquement est de type ETS(ANN), c'est-à-dire un modèle de lissage exponentiel avec un niveau et une tendance, sans saisonnalité explicite. La distribution des erreurs est normale, ce qui est adapté aux séries financières, avec une estimation de l'erreur standard de 0.0463. La trajectoire de prévision reste relativement stable sur l'année 2019, autour d'un niveau de 8.4 à 8.5, avec des fluctuations modérées. Ce modèle est particulièrement efficace pour capter la dynamique à court terme des données, tout en restant adapté aux hétéroscédasticités.

TABLE 7 – Paramètre de lissage α du modèle ADAM ETS

Paramètre	Valeur estimée
α	1.000

Lissage très réactif du niveau (fort poids du présent) Le coefficient $\alpha = 1.000$ indique que le modèle attribue une très forte réactivité aux observations récentes. c'est un lissage très réactif du niveau (fort poids du présent). Cela signifie que chaque nouvelle observation a un impact presque total sur l'ajustement du niveau estimé. Ce comportement est particulièrement adapté aux séries financières telles que le CAC 40, où les variations rapides et immédiates sont fréquentes.

Le modèle ne présente pas de composante saisonnière explicite, ce qui le distingue des

modèles ETS + SARIMA qui incluraient une correction des effets saisonniers. Néanmoins, il reste capable de s'ajuster aux variations de court terme, ce qui en fait un modèle robuste pour des prévisions dans un environnement économique instable (voir annexe 7 et annexe 10).

2.1.4 ADAM ETS + ARIMA

Le modèle ADAM ETS + ARIMA sélectionné automatiquement combine une structure de lissage exponentiel de type ETS(ANN), c'est-à-dire avec niveau et tendance, mais sans composante saisonnière, avec une dynamique SARIMA(0,1,1)[12] appliquée aux résidus. Ce modèle prend en compte les effets saisonniers grâce à une composante de moyenne mobile saisonnière d'ordre 12, tout en capturant des tendances et des chocs non stationnaires avec un paramètre de différenciation d'ordre 1.

Les paramètres estimés sont présentés ci-dessous :

TABLE 8 – Paramètres du modèle ADAM ETS + SARIMA(0,1,1)[12]

Paramètre	Valeur estimée
α	1,1721
θ_{12} (SARIMA)	-1,1399
\hat{a}_0 (niveau initial)	8,4454
\hat{b}_0 (tendance initiale)	-0,0117

Le coefficient $\alpha = 1,1721$ indique une forte réactivité du niveau estimé aux observations récentes (fort poids du présent), chaque nouvelle valeur ayant un grand poids dans la mise à jour du modèle. Ce comportement est particulièrement adapté aux séries financières comme le CAC 40, qui subissent souvent des variations abruptes.

La composante saisonnière est modélisée par un terme de moyenne mobile d'ordre 12 dans la partie SARIMA, avec un coefficient estimé à $\theta_{12} = -1,1399$. Cela signifie qu'il existe une correction des erreurs commises un an auparavant, suggérant une certaine régularité saisonnière, bien que modérée.

Les prévisions issues de ce modèle restent relativement stables sur l'année 2019, avec une légère baisse des valeurs prévues, passant de 8,4454 à 8,4250 à la fin de l'année.

Cette évolution est cohérente avec la pente estimée de $\hat{b}_0 = -0,0117$, ce qui traduit une anticipation d'une légère décroissance dans les valeurs futures. L'erreur standard de prévision est modérée (0,0433) et les critères d'information ($AIC = -769,3345$, $AICc = -766,7563$) indiquent un ajustement satisfaisant (voir annexe 8 et annexe 11).

2.1.5 SSARIMA(0,1,3)(0,0,3)[12]

Tout d'abord, le modèle SSARIMA est une reformulation des modèles ARIMA classiques sous la forme d'un système d'équations d'état. Cette approche, proposée notamment par Svetunkov et Boylan (2020), permet une estimation plus souple et plus performante.

Dans sa forme classique, un modèle ARIMA(0, 1, 3) saisonnier avec une composante de moyenne mobile saisonnière est défini comme suit :

$$(1 - B)(1 - \phi_1 B)y_t = (1 + \theta_1 B + \theta_2 B^2)\varepsilon_t \quad (3)$$

$$\Leftrightarrow y_t = (1 + \phi_1)y_{t-1} - \phi_1 y_{t-2} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t \quad (4)$$

Ce modèle peut être représenté de façon équivalente en forme *état-espace* SSARIMA à l'aide de deux composantes latentes $v_{1,t}$ et $v_{2,t}$, selon :

$$y_t = v_{1,t-1} + v_{2,t-2} + \varepsilon_t \quad (5)$$

$$v_{1,t} = (1 + \phi_1)(v_{1,t-1} + v_{2,t-2}) + (1 + \phi_1 + \theta_1)\varepsilon_t \quad (6)$$

$$v_{2,t} = -\phi_1(v_{1,t-1} + v_{2,t-2}) + (-\phi_1 + \theta_2)\varepsilon_t \quad (7)$$

Ce changement de représentation permet d'interpréter les dynamiques ARIMA comme des interactions entre plusieurs processus latents. Cette formulation facilite l'introduction de plusieurs composantes saisonnières (par exemple à des fréquences différentes comme 12 mois) et rend l'estimation plus stable dans certains contextes, en particulier pour les séries courtes ou irrégulières.

Dans notre cas le modèle SSARIMA(0,1,3)[1](0,0,3)[12] s'est révélé adéquat pour modéliser la dynamique du CAC 40 en log. Ce modèle inclut une différenciation d'ordre 1 pour stabiliser la moyenne de la série, ainsi que trois termes de moyenne mobile non saisonniers (MA). Il intègre également une composante saisonnière avec trois termes de

moyenne mobile saisonnière à retard 12 mois. Il est à noter qu'il ne contient aucun terme autorégressif non saisonnier.

Les coefficients estimés révèlent l'impact des effets saisonniers à travers les termes de moyenne mobile saisonnière MA(1) à 0.0963 et MA(12) à -0.1062 , ainsi que des influences passées, notamment avec les coefficients MA(2) à -0.0932 et MA(3) à 0.1638, traduisant la dynamique des erreurs passées à court et moyen terme.

Le modèle présente une erreur standard de 0.0432, ce qui témoigne d'un bon ajustement aux données. Les critères d'information sont également satisfaisants, avec un AIC de -778.35 et un AICc de -777.84 , confirmant un bon compromis entre précision et prudence.

La prévision à 12 mois générée par ce modèle suit de manière cohérente la tendance récente de la série, avec une légère baisse initiale avant de se stabiliser. Le graphique des prévisions montre des intervalles de confiance étroits à court terme, suggérant une grande fiabilité du modèle dans ses prévisions. De plus, l'ajustement visuel sur la période d'estimation montre des valeurs ajustées proches des observations réelles, ce qui renforce l'idée que le modèle SSARIMA est bien adapté pour la prévision du CAC 40 à court terme (voir annexe 9 et annexe 12).

2.2 Paramètres des modèles précédents et détermination du meilleur modèle d'après les critères AIC et AICc

Pour comparer la performance des différents modèles de prévision, nous utilisons deux critères d'information classiques : l'AIC (Akaike Information Criterion) et sa version corrigée, l'AICc.

L'AIC mesure le compromis entre qualité d'ajustement du modèle et complexité, selon la formule :

$$\text{AIC} = 2k - 2 \ln(\hat{L})$$

où k est le nombre de paramètres estimés et \hat{L} la valeur de la vraisemblance maximisée. Plus l'AIC est faible, meilleur est le compromis.

Cependant, lorsque la taille de l'échantillon n est relativement faible par rapport au nombre de paramètres k , l'AIC peut favoriser des modèles trop complexes. L'AIC corrigé (AICc) introduit un terme supplémentaire :

$$\text{AICc} = \text{AIC} + \frac{2k(k+1)}{n-k-1}$$

Ce critère est donc plus fiable dans les contextes à petits échantillons.

Sur cette base, nous présentons ci-dessous un tableau de synthèse regroupant les principaux paramètres estimés ainsi que les valeurs des critères AIC et AICc pour l'ensemble des modèles testés.

TABLE 9 – Comparaison des paramètres estimés et des critères d'information pour les modèles retenus.

Modèle	ARIMA(1,1,1)	LED Holt-Winters	ADAM ETS	ADAM ETS + ARIMA	SSARIMA
alpha	–	1.000	1.000	1.1721	–
beta	–	0.1795	–	0.000	–
ar1	-0.4494	–	–	0.329	0.0963
ma1	0.6325	–	–	-1.1399	-0.1062
ma2	–	–	–	1.038	-0.0994
ma3	–	–	–	-0.674	0.1638
ma4	–	–	–	–	–
theta	–	–	–	–	–
sar1	–	–	–	–	0.3038
sar2	–	–	–	–	-0.2509
sar3	–	–	–	–	-0.0113
sma1	–	–	–	-1.142	-0.3768
sma2	–	–	–	–	0.1689
sma3	–	–	–	–	0.0080
mean	8.379	–	8.5012	8.4454	–
const	–	–	–	-0.012	–
AIC	-756.35	-1368.59	-751.55	-769.33	-778.35
AICc	-756.24	-1368.53	-751.44	-766.76	-777.84

Par conséquent, en analysant le tableau ci-dessus et en recherchant le modèle présentant les valeurs les plus faibles des critères AIC et AICc, le modèle Holt-Winters (lissage exponentiel double) se distingue clairement, avec un AIC de $-1368,59$ et un AICc de $-1368,53$. Ce modèle, qui capture le niveau et la tendance de la série sans composante saisonnière, est ainsi identifié comme le plus performant parmi ceux testés.

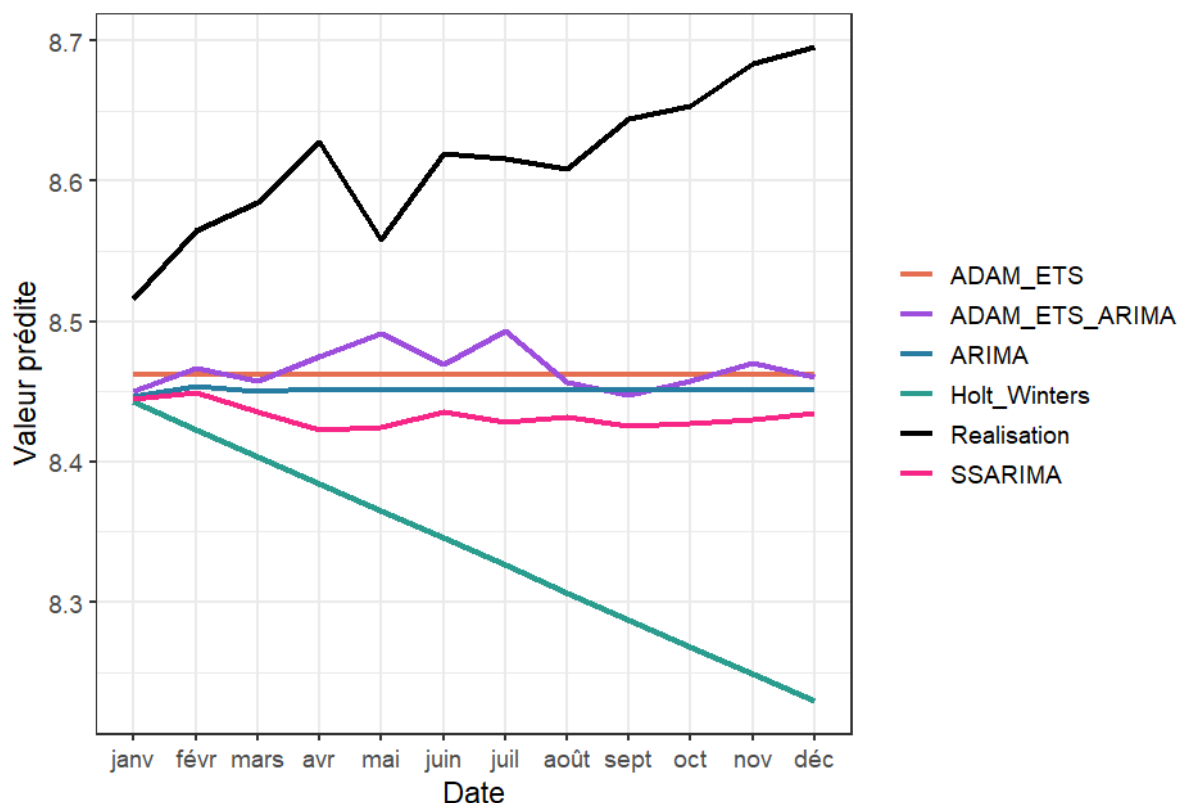
3 Prédiction linéaire sur une année avec représentation graphique

Dans cette partie, nous réalisons une prédiction linéaire du logarithme de l'indice CAC 40 sur une période de 12 mois, correspondant à l'année 2019. L'objectif est de comparer les performances prédictives des différents modèles estimés précédemment sur la base d'apprentissage allant de janvier 2000 à décembre 2018.

Les prévisions sont obtenues sans pas glissant, chaque modèle est estimé une seule fois sur l'ensemble de la période d'apprentissage, puis utilisé pour générer une trajectoire complète sur 12 mois (voir annexes 13, 14, 15, 16 et 17)

Tous les modèles précédents ont été retenus pour cette phase de prédiction. Les trajectoires prédites seront comparées à la série réalisée (valeurs observées de 2019), et serviront de base à l'analyse de la qualité des prévisions dans les sections suivantes (voir annexe 18).

FIGURE 3 – Prévisions mensuelles pour l'année 2019 : comparaison des modèles



Tout d’abord, on remarque que la trajectoire observée suit une tendance haussière tout au long de l’année 2019, indiquant une dynamique positive des marchés financiers dans un contexte de reprise économique mondiale, de stabilité politique en zone euro, et de politiques monétaires accommodantes.

On constate que la série réalisée suit une dynamique globalement haussière, traduisant un climat de marché favorable. Cependant, tous les modèles ne parviennent pas à capturer cette évolution de manière satisfaisante, due au fait qu’en 2018 il y a eu une grande baisse.

Le modèle ADAM ETS+ARIMA se rapproche le plus de la trajectoire réalisée, notamment grâce à sa capacité à intégrer des variations de court terme tout en conservant une bonne stabilité. Sa prévision est cohérente avec le profil réel du marché, avec une légère sous-estimation systématique par rapport aux valeurs observées. Ce modèle est flexible et parvient à suivre efficacement la dynamique haussière du logarithme du CAC 40 sur l’année 2019.

Le modèle ADAM ETS, bien que précis dans l’ensemble, présente une trajectoire constante sur l’année. Cette rigidité traduit une moindre capacité à capter les fluctuations mensuelles du marché, ce qui limite son adaptabilité face aux ajustements progressifs de la dynamique réelle.

Le modèle ARIMA prévoit également une trajectoire relativement stable, mais avec une légère décroissance, ce qui va à l’encontre de la tendance haussière observée en 2019. Cette mauvaise orientation peut être attribuée à une dépendance excessive aux dernières observations de l’échantillon, biaisant les prévisions vers le bas.

Le modèle Holt-Winters affiche la plus forte divergence par rapport à la série réalisée, avec une chute continue et irréaliste sur toute l’année. Ce comportement traduit une mauvaise interprétation des données récentes, avec un biais baissier marqué et sans ajustement aux retournements positifs de marché.

Le modèle SSARIMA, quant à lui, reste relativement stable, bien qu’il sous-estime systématiquement les niveaux atteints par le CAC 40. Sa trajectoire est plus proche de celle de l’ADAM.ETS, mais il présente une flexibilité moindre par rapport au modèle ADAM ETS+ARIMA.

Pour l’analyse graphique, le modèle ADAM ETS+ARIMA apparaît comme le plus

performant dans cet exercice de prévision glissante à court terme. Cette première comparaison visuelle sera consolidée par une analyse quantitative des erreurs de prévision présentée dans les sections suivantes.

4 Qualité de prévision

Dans cette section, nous évaluons la performance des différents modèles à l'aide de critères standards de prévision, notamment l'erreur quadratique moyenne (MSE). Une comparaison visuelle entre les prévisions produites et la série effectivement observée sur une année permettra également d'apprécier la précision de chaque méthode hors échantillon.

4.1 MSE et R_{OOS}^2 des modèles précédents

Pour évaluer la qualité des prévisions produites par les différents modèles, deux indicateurs sont utilisés : l'erreur quadratique moyenne (MSE) et le coefficient de détermination hors-échantillon (R_{OOS}^2).

L'erreur quadratique moyenne (MSE) mesure l'écart moyen au carré entre les valeurs prévues et les valeurs réalisées, elle est définie par la formule suivante :

$$\text{MSE} = \frac{1}{H} \sum_{h=1}^H (y_{t+h} - \hat{y}_{t+h})^2 = \frac{1}{H} \sum_{h=1}^H e_{t+h}^2$$

où y_{t+h} est la valeur observée, \hat{y}_{t+h} la prévision, e_{t+h} l'erreur de prévision, et H l'horizon (ici, 12 mois).

Le coefficient de détermination hors-échantillon (R_{OOS}^2) permet de comparer les performances d'un modèle donné à celles d'un modèle de référence (benchmark), il est défini par la formule suivante :

$$R_{OOS}^2 = 1 - \frac{MSE_i}{MSE_0}$$

où MSE_i est le MSE du modèle évalué, et MSE_0 celui du modèle benchmark (souvent un modèle naïf comme la moyenne ou la dernière valeur observée).

Un coefficient R_{OOS}^2 strictement positif signifie que le modèle considéré fournit de meilleures prévisions que le modèle de référence. En revanche, si R_{OOS}^2 est négatif, cela indique que le modèle évalué est moins performant que le benchmark.

Dans notre étude, nous avons retenu le modèle ARIMA comme benchmark de référence, car il constitue une approche standard et robuste en prévision de séries temporelles financières. Les valeurs du R_{OOS}^2 présentées ci-après permettent donc d'apprécier la performance relative de chaque modèle par rapport à ce benchmark.

TABLE 10 – Comparaison des performances prédictives des modèles (prévision sur l'année 2019, MSE et R_{OOS}^2 , benchmark ARIMA)

Modèle	MSE	R_{OOS}^2
ADAM ETS.ARIMA	0,0247	0,152
ADAM ETS	0,0258	0,115
ARIMA	0,0291	–
Holt-Winters	0,0906	-2,11
SSARIMA	0,0360	-0,237

Les résultats présentés dans le tableau ci-dessus montrent une variation significative des performances des différents modèles évalués, avec des résultats particulièrement intéressants pour certains modèles par rapport à l'ARIMA (voir annexe 19).

Le modèle ADAM ETS+ARIMA, qui combine le lissage exponentiel et la structure ARIMA, affiche un MSE légèrement inférieur (0,0247) au modèle ADAM ETS, ainsi qu'un R_{OOS}^2 de 0,152 supérieur. Ce résultat montre que le modèle parvient à réduire l'erreur de prévision de plus de 15 % par rapport au modèle de référence ARIMA, confirmant ainsi sa robustesse et sa bonne capacité d'adaptation aux dynamiques de marché à court terme.

Le modèle ADAM ETS se distingue également des autres avec une MSE de 0,0258 et un R_{OOS}^2 de 0,115. Il présente une précision brute correcte, bien qu'inférieure à celle du modèle ADAM ETS+ARIMA, comme en témoignent une erreur quadratique moyenne plus élevée et une amélioration hors échantillon moindre par rapport à ARIMA. Il apporte ainsi un gain significatif en performance prédictive, même si celui-ci reste inférieur à celui offert par le modèle ADAM ETS+ARIMA.

Le modèle SSARIMA présente un MSE de 0,0360 et un R_{OOS}^2 de -0,237, ce qui traduit une performance inférieure à celle d'ARIMA. Malgré sa relative stabilité, il ne parvient pas à capter efficacement les dynamiques sous-jacentes de la série, entraînant une perte de précision notable.

Le modèle Holt-Winters enregistre de loin la pire performance, avec un MSE de 0,0906 et un R_{OOS}^2 de -2,11. Cette très forte sous-performance traduit l'incapacité du modèle à suivre la tendance réelle du logarithme du CAC 40, principalement en raison d'une mauvaise extrapolation d'une tendance baissière inexistante sur la période.

En résumé, le modèle ADAM ETS+ARIMA apparaît comme le plus adapté pour la prévision glissante sur l'année 2019, surpassant clairement ARIMA en termes d'erreur quadratique et de capacité de généralisation hors-échantillon. ADAM ETS montre également de bonnes performances, mais avec un gain limité en comparaison.

4.2 Comparaison avec une prévision naïve

Afin d'évaluer la performance des modèles testés, nous allons utiliser une prévision dite naïve comme point de comparaison. Cette prévision repose sur une règle simple, qui est la supposition que la meilleure anticipation de la valeur future est simplement la dernière valeur observée. Elle sert alors de *benchmark* de référence.

La prévision naïve est définie par :

$$\hat{y}_{t+h|t}^{\text{naïve}} = y_t \quad \text{pour tout } h \geq 1$$

Nous avons donc comparé les performances de chaque modèle à cette référence naïve en recalculant le R_{OOS}^2 avec MSE_0 correspondant à l'erreur quadratique moyenne de cette prévision. Cette analyse permet de confirmer si les modèles testés apportent une réelle amélioration par rapport à une stratégie minimale d'anticipation.

Dans cette comparaison, la prévision naïve retenue repose sur le modèle naïf avec persistance, estimé à l'aide de la fonction `naive()` sur la série log-transformée du CAC 40. Cette approche consiste à répliquer la dernière valeur observée de la période d'apprentissage (décembre 2018) pour chaque mois de l'année 2019. La série de prévisions obtenue est donc constante, et sert de benchmark pour le calcul du $R_{OOS, \text{naïve}}^2$ des autres modèles.

TABLE 11 – Comparaison des performances prédictives des modèles avec la prévision naïve et ARIMA comme benchmark

Modèle	MSE	R_{OOS}^2
ADAM ETS+ARIMA	0,0247	0,0421
ADAM ETS	0,0258	0
Naïve	0,0258	–
ARIMA	0,0291	-0,129
SSARIMA	0,0360	-0,397
Holt-Winters	0,0906	-2,51

L’analyse des résultats présentés dans le tableau met en évidence la difficulté des modèles testés à surpasser la prévision naïve sur l’année 2019 (voir annexe 19, 20 et 21).

Le modèle ADAM ETS+ARIMA présente un R_{OOS}^2 de 0,0421, indiquant une légère amélioration par rapport à la naïve (environ 4,2%). Cette performance montre que la combinaison des approches ETS et ARIMA permet d’intégrer un peu plus d’information structurelle, mais sans offrir un gain substantiel dans ce contexte précis.

Le modèle ADAM ETS obtient un MSE identique à celui de la prévision naïve (0,0258) et un R_{OOS}^2 de 0. Ce résultat s’explique par la structure même du modèle. En effet, en l’absence de tendance ou de saisonnalité détectée sur la série, ADAM ETS a sélectionné une forme très simple (type ETS(A,N,N)), équivalente à une prévision naïve répétant simplement la dernière valeur observée. Ainsi, le modèle reproduit exactement les performances de la stratégie naïve sans amélioration ni dégradation.

Le modèle ARIMA, avec un R_{OOS}^2 de -0,129, montre une sous-performance face à la stratégie naïve. Cela suggère que ce modèle quasiment linéaire n’a pas réussi à saisir les subtilités de la dynamique boursière sur la période étudiée.

Le modèle SSARIMA affiche un R_{OOS}^2 de -0,397, ce qui indique également qu’il sous-performe significativement par rapport à la prévision naïve. Ce modèle est conçu pour mieux traiter les dynamiques saisonnières, mais celles-ci ne sont pas suffisamment marquées ou correctement capturées sur la période estimée, étant donné la non-saisonnalité de la série du CAC 40.

Enfin, le modèle de Holt-Winters présente la plus mauvaise performance, avec un R_{OOS}^2 de -2,51. Cette très forte dégradation par rapport à la prévision naïve souligne l'inadéquation de ce modèle, qui extrapole une tendance baissière inexistante et échoue à suivre les évolutions réelles du logarithme du CAC 40 (et donc de son niveau également).

En résumé, le modèle ADAM ETS+ARIMA apparaît de nouveau comme le plus pertinent.

4.3 Graphique des CSPE

Après avoir comparé les performances globales des modèles à l'aide du MSE et du R_{OOS}^2 , par rapport au modèle ARIMA et à une prévision naïve, il est également pertinent d'examiner la stabilité des erreurs de prévision dans le temps.

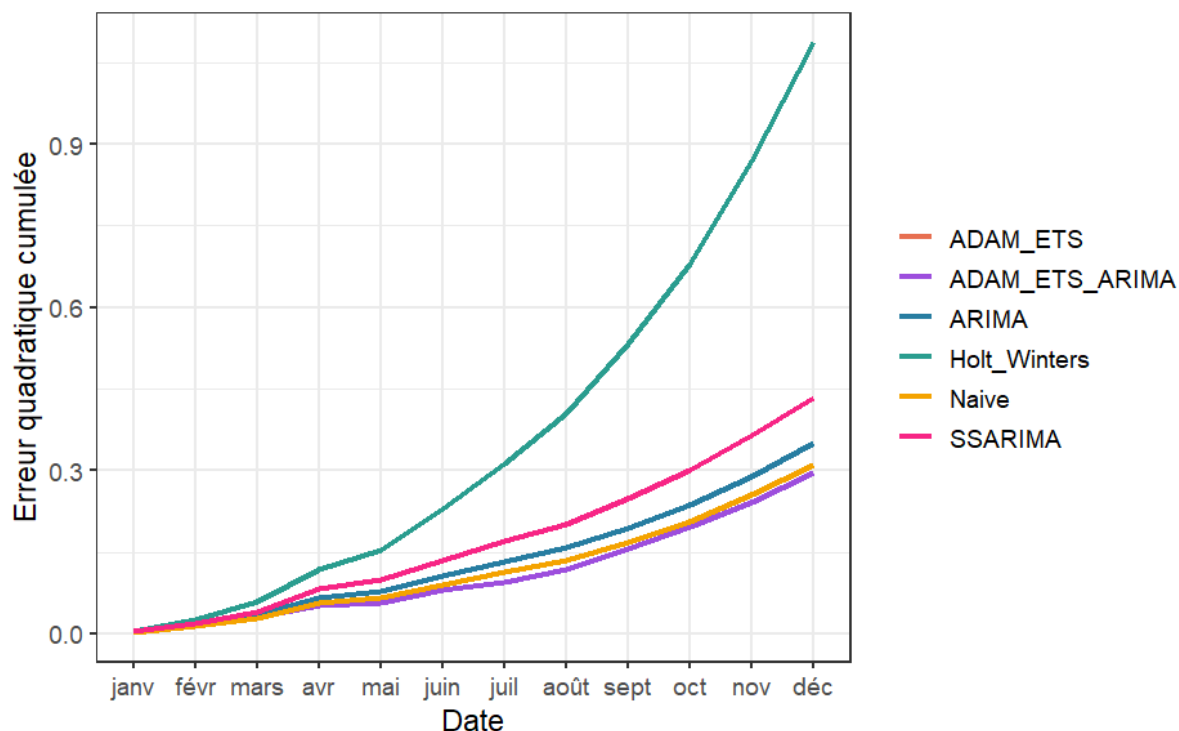
Pour cela, nous utilisons le graphique des erreurs cumulées de prévision au carré (*Cumulative Squared Prediction Error*, CSPE). Ce graphique permet de visualiser, pour chaque modèle, l'accumulation progressive des erreurs quadratiques de prévision sur l'horizon considéré.

La formule du CSPE est donnée par :

$$CSPE = \sum_{h=1}^H (y_{t+h} - \hat{y}_{t+h})^2 = \sum_{h=1}^H e_{t+h}^2$$

Un modèle efficace se caractérise par une croissance plus lente de la courbe CSPE, indiquant une meilleure précision de prévision tout au long de la période. À l'inverse, une courbe qui croît rapidement traduit une accumulation importante d'erreurs, donc une faible stabilité prédictive.

FIGURE 4 – Évolution cumulée de l’erreur quadratique de prévision (log-CAC 40, année 2019)



On observe que le modèle ADAM ETS+ARIMA présente des erreurs quadratiques légèrement inférieures à celles du modèle ADAM ETS seul ainsi qu’à celles de la prévision naïve (ces deux dernières ayant les mêmes valeurs). Bien que l’amélioration ne soit pas spectaculaire, elle montre que ce modèle parvient à mieux ajuster ses prévisions, en particulier dans des contextes où les tendances sont plus complexes.

Comme précédemment, le modèle ADAM ETS se distingue par le fait qu’il se superpose à la courbe de l’erreur quadratique de prévision cumulée du modèle naïf. Cela s’explique par la structure simple du modèle, équivalente à celle de la prévision naïve (voir annexe 22).

Ensuite, les courbes CSPE des modèles ARIMA et SSARIMA montrent une accumulation des erreurs au fil du temps, supérieure à celle des trois premiers modèles. Dès le début de la période de prévision, ces courbes présentent une pente plus prononcée, ce qui suggère que ces modèles ont davantage de difficultés à s’adapter rapidement aux variations de la série.

Enfin, la courbe du modèle Holt-Winters présente une courbe CSPE fortement croissante dès les premiers mois, indiquant une accumulation rapide d'erreurs. Cette instabilité suggère que le modèle ne parvient pas à s'ajuster efficacement aux évolutions récentes du CAC 40. Cela reflète la difficulté du modèle à capturer correctement la dynamique du marché et son manque de flexibilité face aux variations.

5 Test de précision

Dans cette section, nous cherchons à comparer la précision prédictive des différents modèles estimés face à une prévision naïve, en mobilisant le test de Diebold-Mariano.

5.1 Test de Diebold-Mariano

Le test de Diebold-Mariano permet de comparer la précision de deux méthodes de prévision en testant si leurs erreurs prévisionnelles sont significativement différentes. Il repose sur une fonction de perte $g(\cdot)$, comme la fonction de l'erreur quadratique moyenne (MSE), et utilise la différence des pertes :

$$d_t = g(e_t^{(1)}) - g(e_t^{(2)})$$

où $e_t^{(1)}$ et $e_t^{(2)}$ sont les erreurs de prévision des deux modèles comparés à l'instant t . Si les deux modèles ont une précision similaire, alors l'espérance de d_t est nulle : $E[d_t] = 0$.

Le test repose alors sur la statistique :

$$DM = \frac{\bar{d}}{\sqrt{\hat{V}(\bar{d})}} \quad \text{où} \quad \bar{d} = \frac{1}{T} \sum_{t=1}^T d_t$$

avec $\hat{V}(\bar{d})$ une estimation de la variance de \bar{d} , généralement obtenue avec un estimateur de Newey-West (notamment si l'horizon de prévision $h > 1$).

Dans le cadre du test de Diebold-Mariano, les hypothèses sont formulées en termes de l'espérance de la différence des pertes d_t entre deux modèles. L'hypothèse nulle est $H_0 : E[d_t] = 0$, ce qui signifie que les deux modèles ont des précisions équivalentes en moyenne.

L'hypothèse alternative dépend de la direction du test : si l'on suppose que le modèle 1 est moins précis que le modèle 2, on teste $H_1 : E[d_t] > 0$; à l'inverse, si l'on souhaite vérifier si le modèle 1 est plus précis que le modèle 2, on utilise $H_1 : E[d_t] < 0$.

Dans notre étude, nous utilisons l'option "less" dans le test, afin de vérifier si chaque modèle prédit significativement mieux que la prévision naïve.

TABLE 12 – Test de Diebold-Mariano : comparaison de chaque modèle à la prévision naïve (hypothèse alternative : « less »)

Modèle	p-value (DM test vs naïve)
ADAM ETS+ARIMA	0,828
ADAM ETS	NA
Holt-Winters	0,000643
SSARIMA	0,00000396
ARIMA	0,000000133

Les résultats du test de Diebold-Mariano (voir annexe 23), effectués pour chaque modèle en comparaison avec la prévision naïve, indiquent, de par leur p-value inférieure au seuil de 1 %, que les modèles Holt-Winters, SSARIMA et ARIMA rejettent l'hypothèse nulle au seuil de 1 %. Par conséquent, on rejette l'égalité des prévisions et les prévisions issues de ces modèles sont significativement différentes et inférieures à celles de la prévision naïve.

Comme précédemment mentionné, le résultat du test pour le modèle ADAM ETS est indiqué comme NA, en raison de la parfaite similarité de ses prévisions avec celles du modèle naïf, rendant la comparaison inapplicable.

Concernant le modèle ADAM ETS+ARIMA, la p-value du test est très largement supérieure à 5%, ce qui conduit à ne pas rejeter l'hypothèse nulle. Il n'existe donc pas de preuve statistique que les prévisions du modèle ADAM ETS+ARIMA soient significativement différentes de celles de la prévision naïve. Les deux méthodes offrent des performances de prévision considérées comme équivalentes dans ce cadre.

5.2 Définition du meilleur modèle

Parmi l'ensemble des modèles évalués, le modèle ADAM ETS+ARIMA ressort comme le plus performant en matière de prévision hors-échantillon. Il présente la plus faible erreur quadratique moyenne parmi tous les modèles testés, traduisant une meilleure précision globale de prévision par rapport aux autres approches.

Toutefois, le test de Diebold-Mariano indique que la différence de performance entre

ADAM ETS+ARIMA et la prévision naïve n'est pas statistiquement significative au seuil de 5%, la p-value étant largement supérieure à ce seuil. Cela signifie que, bien que le modèle ADAM ETS+ARIMA améliore légèrement la précision par rapport à la naïve, cette amélioration reste modeste d'un point de vue strictement statistique.

Ces résultats suggèrent néanmoins que le modèle ADAM ETS+ARIMA parvient à capter plus efficacement les dynamiques de la série étudiée, en offrant des prévisions plus stables et plus précises. Il constitue donc une option robuste pour la prévision à court terme du CAC 40.

6 Prédiction sur une année avec un pas de 1 mois

Afin d'évaluer la robustesse du modèle sélectionné, nous réalisons une prédiction glissante sur l'année 2019 à un horizon d'un mois. Le principe consiste à estimer le modèle ADAM ETS+ARIMA à chaque mois t en utilisant uniquement les données disponibles jusqu'à $t - 1$, puis à produire une prédiction pour le mois t . Ce processus est répété 12 fois, une fois pour chaque mois de l'année.

Le tableau ci-dessous présente, pour chaque mois de 2019, la valeur réalisée du CAC 40 (transformée en logarithme), la prédiction obtenue par le modèle ADAM ETS+ARIMA, ainsi que les erreurs associées (brute et au carré) :

TABLE 13 – Prédiction glissante mensuelle sur l'année 2019 — Modèle ADAM.ETS+ARIMA (log)

Mois	Réalisation	Prédiction (ADAM ETS+ARIMA)	Erreur	Erreur ²
Janvier	8.52	8.45	0.07	0.0049
Février	8.56	8.53	0.03	0.0009
Mars	8.58	8.57	0.01	0.0001
Avril	8.63	8.60	0.03	0.0009
Mai	8.56	8.64	-0.08	0.0064
Juin	8.62	8.53	0.09	0.0081
Juillet	8.62	8.65	-0.03	0.0009
Août	8.61	8.60	0.01	0.0001
Septembre	8.64	8.61	0.03	0.0009
Octobre	8.65	8.65	0.00	0.0000
Novembre	8.68	8.66	0.02	0.0004
Décembre	8.70	8.66	0.04	0.0016

Les erreurs de prédiction du modèle ADAM ETS+ARIMA demeurent globalement faibles sur l'ensemble de l'année 2019. Le modèle démontre une excellente précision, notamment en février, mars, juillet et août, où l'écart par rapport aux valeurs réalisées reste inférieur à 0,02. Cette performance traduit une bonne capacité d'ajustement face aux évolutions modérées du marché (voir annexe 24).

En revanche, des écarts plus marqués sont observés en mai et juin (avec des erreurs

proches ou supérieures à 0,07), ainsi qu'en novembre et décembre, mettant en évidence certaines limites du modèle lorsqu'il est confronté à des retournements de tendance ou à des dynamiques de fin d'année plus instables.

Malgré ces fluctuations ponctuelles, le modèle parvient à suivre avec précision la trajectoire générale du CAC 40 (log-transformé), comme en atteste la forte proximité des courbes de prévision et de réalisation sur la majorité des mois (voir annexe 24). Cette performance souligne la robustesse du modèle ADAM ETS+ARIMA dans un exercice de prévision glissante à court terme, bien qu'une certaine sensibilité aux épisodes de volatilité reste perceptible, phénomène classique sur les marchés financiers.

7 Conclusion

Ce dossier avait pour objectif d’analyser, modéliser et prévoir l’évolution de l’indice CAC 40 sur la période 2000–2019, en mobilisant différentes techniques de séries temporelles adaptées aux séries non saisonnières.

L’analyse préliminaire a permis de mettre en évidence plusieurs caractéristiques structurantes de la série étudiée : non-stationnarité en niveau, existence de points atypiques majeurs liés à des événements économiques mondiaux (bulle Internet, crise des subprimes, crise des dettes souveraines) et schéma multiplicatif.

Dans un second temps, nous avons estimé plusieurs modèles, ARIMA, Holt-Winters LED, ADAM ETS, ADAM ETS+ARIMA et SSARIMA. Chaque approche a été évaluée à la fois sur la qualité d’ajustement (AIC, AICc) et sur la capacité prédictive hors-échantillon (MSE, R^2_{OOS} , CSPE). La prévision naïve a également été utilisée comme benchmark pour tester la robustesse des modèles.

Parmi l’ensemble des modèles évalués, le modèle ADAM.ETS+ARIMA est ressorti comme le plus performant en matière de prévision hors-échantillon. Il présente la plus faible erreur quadratique moyenne sur l’année 2019 et offre une trajectoire de prévision plus ou moins similaire à la dynamique observée du CAC 40. Bien que le test de Diebold-Mariano n’ait pas confirmé une amélioration significative au sens strict par rapport à la prévision naïve, les résultats obtenus montrent néanmoins une capacité supérieure d’ajustement et de stabilité prédictive. À noter également, que dans notre cas la prédiction naïve était très bonne.

La prévision glissante réalisée sur l’année 2019 a confirmé la robustesse du modèle ADAM ETS+ARIMA, malgré quelques erreurs liées aux retournements de tendance. Le modèle a su s’adapter aux évolutions du marché, en captant la tendance haussière modérée du CAC 40, tout en conservant une bonne précision de prévision sur des horizons courts.

En définitive, cette étude met en avant l’intérêt d’utiliser des approches combinant le lissage exponentiel et les structures ARIMA pour la prévision de séries financières non saisonnières. Le modèle ADAM ETS+ARIMA apparaît ainsi comme une solution robuste, équilibrant flexibilité et stabilité, adaptée à la prévision conjoncturelle à court terme d’indices boursiers.

8 Bibliographie

Références

- [1] Yahoo Finance. (2025). *Finance news, data, and tools*. Consulté en avril 2025.
- [2] De Jonghe, A. (2017). *La bulle Internet : origines, mécanismes, conséquences*. Disponible en ligne : <https://www.dejonghe-finance.com/wp-content/uploads/2017/07/la-bulle-internet.pdf>, consulté en avril 2025.
- [3] Capital. (2013). *Panorama de la Bourse de Paris, 10 ans après la fin de la bulle Internet*. Disponible en ligne : <https://www.capital.fr/entreprises-marches/panorama-de-la-bourse-de-paris-10-ans-apres-la-fin-de-la-bulle-internet-819005>, consulté en avril 2025.
- [4] Le Figaro. (2009). *CAC 40, d'un 9 mars à un autre...* Disponible en ligne : <https://www.lefigaro.fr/bourse/2009/12/31/04013-20091231ARTFIG00477-cac-40-d-un-9-mars-a-un-autre8230-.php>, consulté en avril 2025.
- [5] Les Échos. (2010). *La zone euro secouée par une grave crise de la dette*. Disponible en ligne : <https://www.lesechos.fr/2010/12/la-zone-euro-secouee-par-une-grave-crise-de-la-dette-437815>, consulté en avril 2025.
- [6] Les Échos. (2013). *Mario Draghi démène le terrain*. Disponible en ligne : <https://www.lesechos.fr/2013/03/mario-draghi-demene-le-terrain-318596>, consulté en avril 2025.
- [7] Le Monde. (2017). *En 2017, les marchés financiers ont retrouvé le goût du risque*. Disponible en ligne : https://www.lemonde.fr/economie/article/2017/12/30/en-2017-les-marches-financiers-ont-retrouve-le-gout-du-risque_5235985_3234.html, consulté en avril 2025.
- [8] Le Parisien. (2019). *Bourse : le CAC 40 termine 2018 en baisse de près de 11%*. Disponible en ligne : <https://www.leparisien.fr/economie/votre-argent/bourse-le-cac-40-termine-2018-en-baisse-de-pres-de-11-01-01-2019-7979010.php>, consulté en avril 2025.

- [9] Les Échos. (2001, septembre). *Le CAC 40 terrassé par les attentats*. Disponible en ligne : <https://www.lesechos.fr/2001/09/le-cac-40-terrasse-par-les-attentats-702827>, consulté en avril 2025.
- [10] Wikipédia. *Scandale Enron*. Disponible en ligne : https://fr.wikipedia.org/wiki/Scandale_Enron, consulté en avril 2025.
- [11] Wikipédia. *Scandale financier WorldCom de 2002*. Disponible en ligne : https://fr.wikipedia.org/wiki/Scandale_financier_WorldCom_de_2002, consulté en avril 2025.
- [12] Le Figaro. (21 janvier 2008). *La Bourse de Paris dévisse sous les points*. Disponible en ligne : <https://www.lefigaro.fr/marches/2008/01/21/04003-20080121ARTFIG00433-la-bourse-de-paris-devisse-sous-les-points.php>, consulté en avril 2025.

9 Annexes

Annexe 1 : Détection des points atypiques et détermination du caractère multiplicatif du schéma.

```
y = regression model + arima (0, 1, 1, 0, 0, 0)

Model: RegARIMA - X13
Estimation span: from 1-2000 to 12-2018
Log-transformation: yes
Regression model: no mean, no trading days effect, no leap year effect, Easter effect, outliers(3)

Coefficients:
ARIMA:
      Estimate Std. Error T-stat Pr(>|t|)
Theta(1)  0.1803    0.0658   2.74 0.00663 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Regression model:
      Estimate Std. Error T-stat Pr(>|t|)
Easter [15]  0.017682    0.008937   1.978 0.04909 *
AO (9-2002) -0.153570    0.029455  -5.214 4.16e-07 ***
TC (9-2001) -0.130310    0.040755  -3.197 0.00158 **
LS (1-2008) -0.139890    0.045255  -3.091 0.00224 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0455 on 221 degrees of freedom
Log likelihood = 379.3, aic = 3051, aicc = 3052, bic(corrected for length) = -6.061
```

Annexe 2 : Stationnarité : test de Dickey-Fuller augmenté.

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression drift

Call:
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-0.12876 -0.02718  0.00566  0.02650  0.10475

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.25900    0.12162   2.130  0.0343 *
z.lag.1      -0.03124    0.01463  -2.135  0.0338 *
z.diff.lag    0.15186    0.06575   2.310  0.0218 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.04518 on 223 degrees of freedom
Multiple R-squared:  0.03835, Adjusted R-squared:  0.02973
F-statistic: 4.447 on 2 and 223 DF, p-value: 0.01277

Value of test-statistic is: -2.1351 2.298

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.46 -2.88 -2.57
phi1  6.52  4.63  3.81
```

Annexe 3 : Stationnarité : test de Philips et Perron.

```
#####
# Phillips-Perron Unit Root Test #
#####

Test regression with intercept

Call:
lm(formula = y ~ y.l1)

Residuals:
    Min       1Q   Median       3Q      Max
-0.145241 -0.024667  0.006403  0.029638  0.100973

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.21203    0.12269   1.728  0.0853 .
y.l1         0.97446    0.01476  66.019 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.04595 on 225 degrees of freedom
Multiple R-squared:  0.9509,    Adjusted R-squared:  0.9507
F-statistic: 4358 on 1 and 225 DF, p-value: < 2.2e-16

Value of test-statistic, type: Z-tau is: -2.1828

      aux. Z statistics
Z-tau-mu      2.181

Critical values for Z statistics:
      1pct      5pct      10pct
critical values -3.460495 -2.874324 -2.573522
```

Annexe 4 : Stationnarité : test de kpss.

```
#####
# KPSS Unit Root Test #
#####

Test is of type: mu with 14 lags.

Value of test-statistic is: 0.3004

Critical value for a significance level of:
      10pct  5pct  2.5pct  1pct
critical values 0.347 0.463 0.574 0.739
```

Annexe 5 : Résultat du code pour le modèle ARIMA.

```
Series: log_cac40_ts
ARIMA(1,1,1)

Coefficients:
      ar1      ma1
    -0.4494  0.6325
s.e.    0.1792  0.1521

sigma^2 = 0.002055: log likelihood = 381.17
AIC=-756.35 AICc=-756.24 BIC=-746.07

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set -0.0001591391 0.0450317 0.03508933 -0.003180929 0.4243453 0.2177952 -0.01436514
```

Annexe 6 : Résultat du code pour le modèle Holt-Winters.

```
Holt-Winters exponential smoothing with trend and without seasonal component.

Call:
HoltWinters(x = log_cac40_ts, gamma = FALSE)

Smoothing parameters:
alpha: 1
beta : 0.1795334
gamma: FALSE

Coefficients:
      [,1]
a 8.46182634
b -0.01936874
```

Annexe 7 : Résultat du code pour le modèle ADAM.ETS(ANN).

```
Model estimated using auto.adam() function: ETS(ANN)
Response variable: data
Distribution used in the estimation: Normal
Loss function type: likelihood; Loss function value: -378.7726
Coefficients:
      Estimate Std. Error Lower 2.5% Upper 97.5%
alpha    1.0000     0.0739     0.8544     1.0000 *
level    8.5012     0.0464     8.4097     8.5927 *

Error standard deviation: 0.0463
Sample size: 228
Number of estimated parameters: 3
Number of degrees of freedom: 225
Information criteria:
      AIC      AICc      BIC      BICc
-751.5452 -751.4381 -741.2572 -740.9663
```

Annexe 8 : Résultat du code pour le modèle ADAM ETS(ANN)+SARIMA.

```
> summary(modele_adam_ets_arima)

Model estimated using auto.adam() function: ETS(ANN)+SARIMA(0,1,1)[12]
Response variable: log_cac40_ts
Distribution used in the estimation: Normal
Loss function type: likelihood; Loss function value: -400.6672
Coefficients:
      Estimate Std. Error Lower 2.5% Upper 97.5%
alpha         1.1721    0.0825    1.0096    1.3346 *
theta1[12]    -1.1399    0.0194   -1.1782   -1.1016 *
level         8.4454    0.0275    8.3913    8.4995 *
ARIMASate1    -0.0117    0.0028   -0.0172   -0.0061 *
ARIMASate2    -0.0002    0.0022   -0.0045    0.0040
ARIMASate3    -0.0061    0.0019   -0.0099   -0.0022 *
ARIMASate4    -0.0332    0.0019   -0.0369   -0.0295 *
ARIMASate5    -0.0229    0.0018   -0.0264   -0.0193 *
ARIMASate6    -0.0022    0.0019   -0.0060    0.0016
ARIMASate7    -0.0005    0.0019   -0.0043    0.0034
ARIMASate8    -0.0080    0.0021   -0.0122   -0.0039 *
ARIMASate9     0.0011    0.0022   -0.0031    0.0054
ARIMASate10   -0.0107    0.0026   -0.0158   -0.0056 *
ARIMASate11    0.0011    0.0030   -0.0049    0.0071
ARIMASate12   -0.0077    0.0031   -0.0139   -0.0016 *

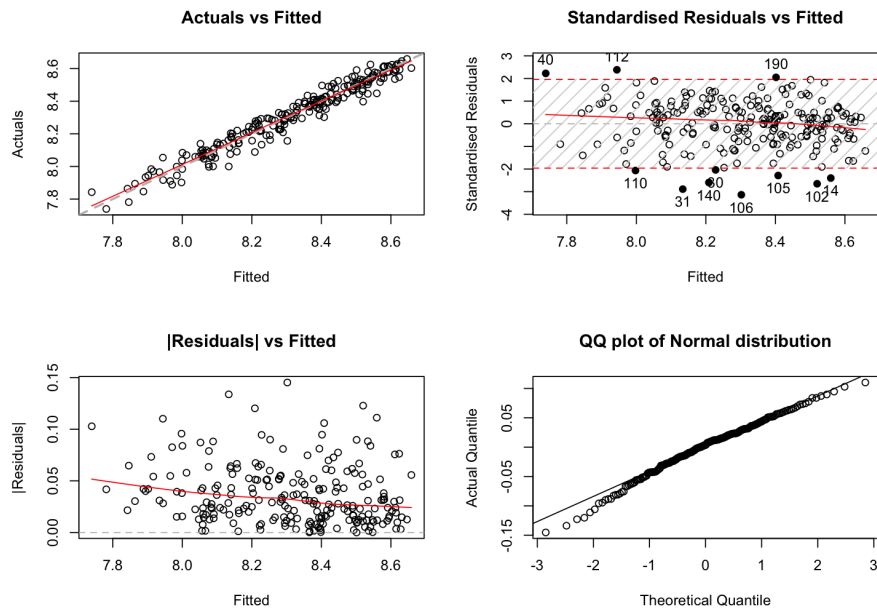
Error standard deviation: 0.0433
Sample size: 228
Number of estimated parameters: 16
Number of degrees of freedom: 212
Information criteria:
      AIC      AICc      BIC      BICc
-769.3345 -766.7563 -714.4650 -707.4660
```

Annexe 9 : Résultat du code pour le modèle SSARIMA.

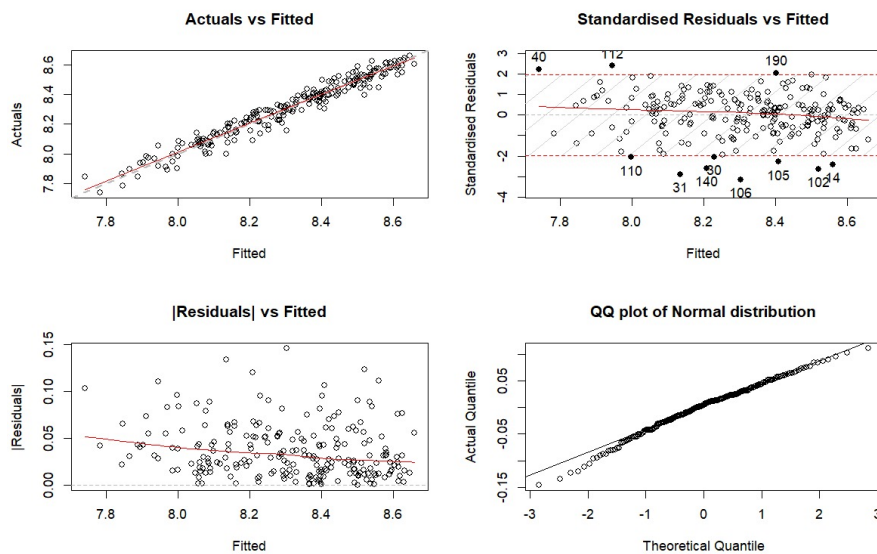
```
Time elapsed: 0.76 seconds
Model estimated: SARIMA(0,1,3)[1](0,0,3)[12]
Matrix of MA terms:
      Lag 1  Lag 12
MA(1)  0.0963 -0.1062
MA(2) -0.0932 -0.0994
MA(3)  0.1638 -0.0426
Initial values were produced using backcasting.

Loss function type: likelihood; Loss function value: -396.1734
Error standard deviation: 0.0432
Sample size: 228
Number of estimated parameters: 7
Number of degrees of freedom: 221
Information criteria:
      AIC      AICc      BIC      BICc
-778.3469 -777.8378 -754.3415 -752.9595
```

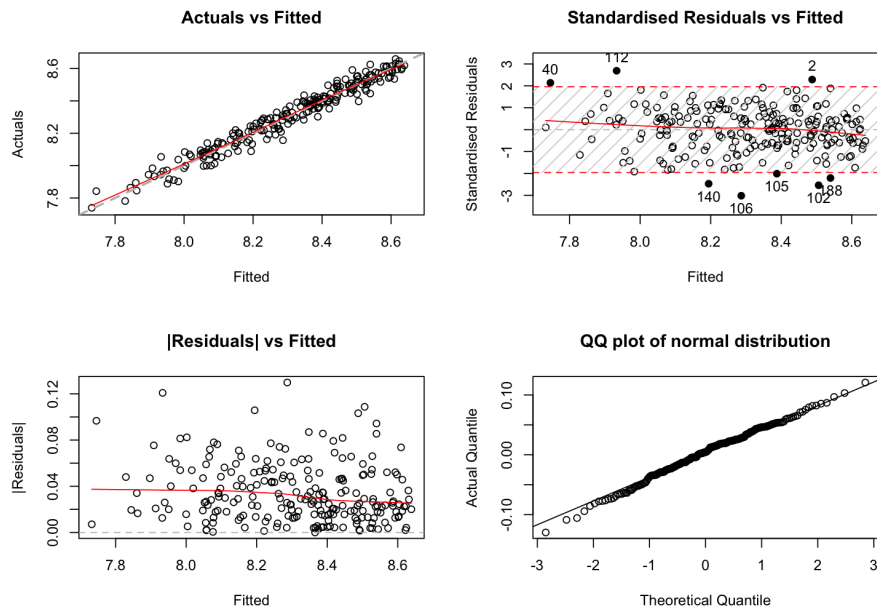
Annexe 10 : Diagnostics des résidus pour le modèle ADAM ETS.



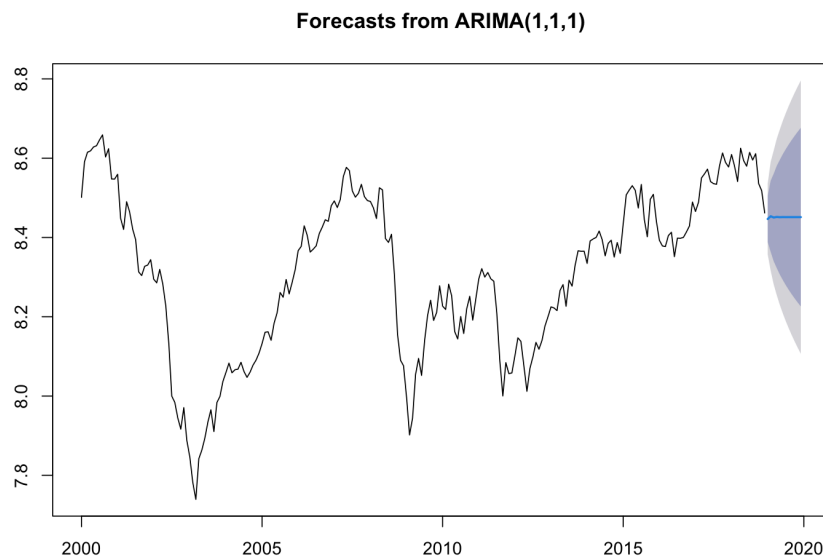
Annexe 11 : Diagnostics des résidus pour le modèle ADAM ETS+ARIMA.



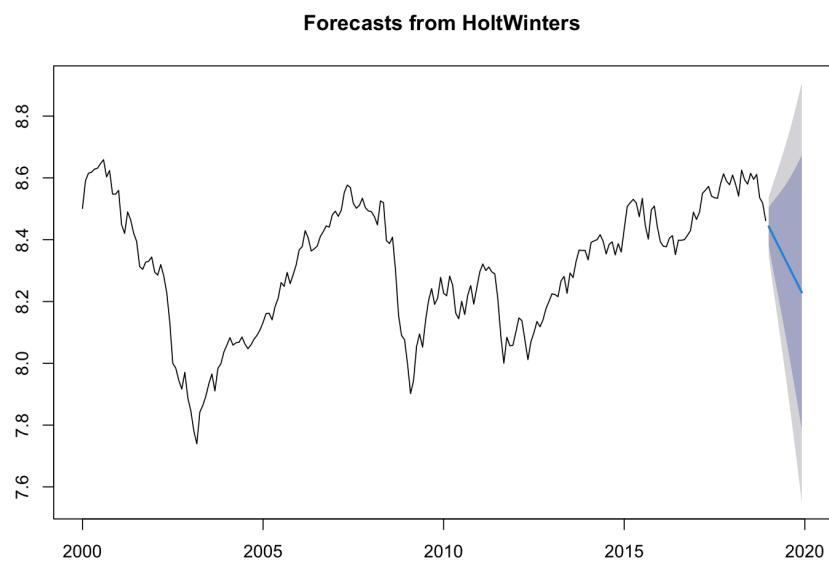
Annexe 12 : Diagnostics des résidus pour le modèle SSARIMA.



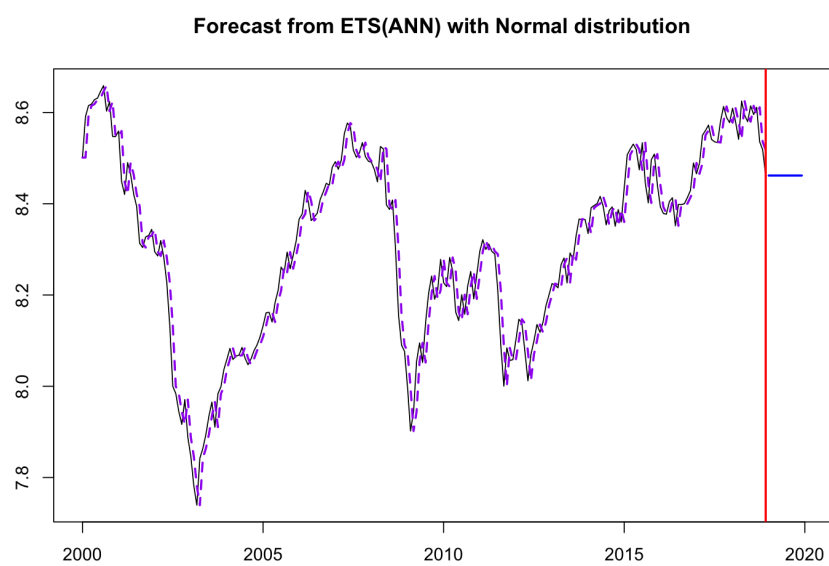
Annexe 13 : Prévisions du modèle ARIMA.



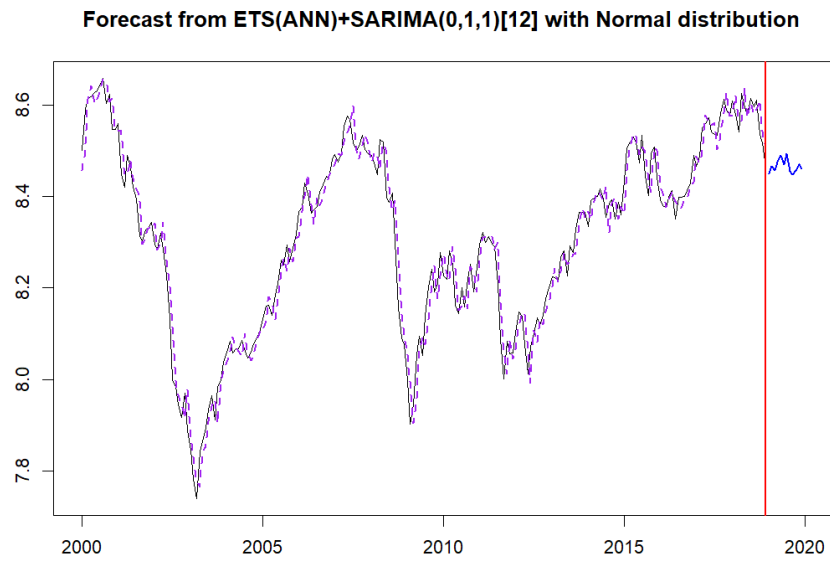
Annexe 14 : Prévisions du modèle Holt-Winters.



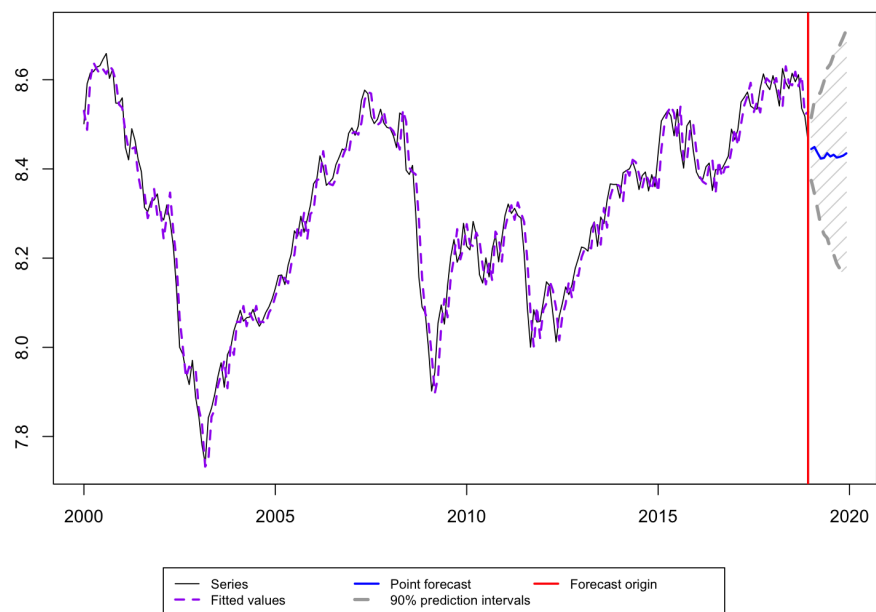
Annexe 15 : Prévisions du modèle ADAM ETS.



Annexe 16 : Prévisions du modèle ADAM ETS+SARIMA.



Annexe 17 : Prévisions du modèle SSARIMA.



Annexe 18 : Prévisions mensuelles de 2019 (en log) par modèle : écart à la réalisation.

Mois	ARIMA	Holt-Winters	ADAM.ETS	ADAM.ETS.ARIMA	SSARIMA	Réalisation
Janvier	8.45	8.44	8.46	8.45	8.44	8.52
Février	8.45	8.42	8.46	8.47	8.45	8.56
Mars	8.45	8.40	8.46	8.46	8.44	8.58
Avril	8.45	8.38	8.46	8.48	8.42	8.63
Mai	8.45	8.36	8.46	8.46	8.42	8.56
Juin	8.45	8.35	8.46	8.47	8.44	8.62
Juillet	8.45	8.33	8.46	8.49	8.43	8.62
Août	8.45	8.31	8.46	8.50	8.43	8.61
Septembre	8.45	8.29	8.46	8.45	8.43	8.64
Octobre	8.45	8.27	8.46	8.46	8.43	8.65
Novembre	8.45	8.25	8.46	8.47	8.43	8.68
Décembre	8.45	8.23	8.46	8.46	8.43	8.70

Annexe 19 : Résultats MSE et R^2_{OOS} .

```
# A tibble: 3 x 7
```

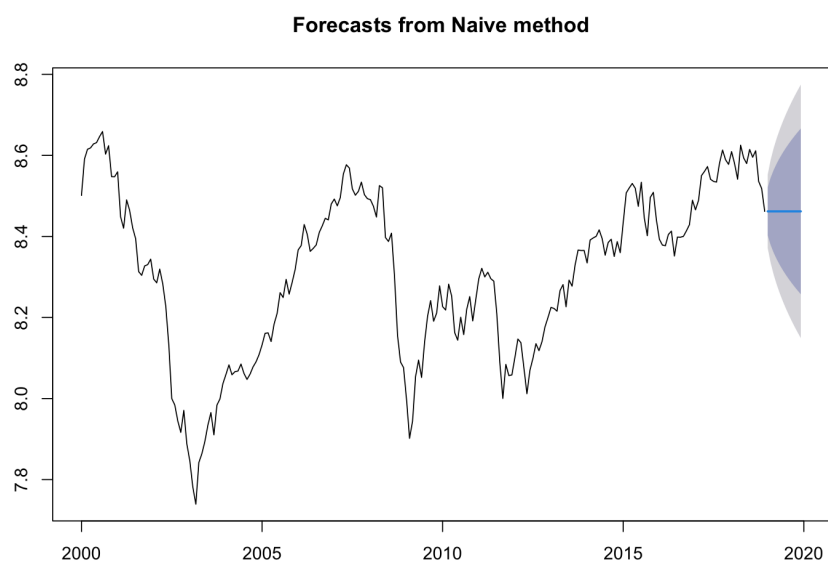
	Critere	Naive	ARIMA	Holt_winters	ADAM_ETS	ADAM_ETS_ARIMA	SSARIMA
	<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	MSE	0.0258	0.0291	0.0906	0.0258	0.0247	0.0360
2	R2_OOS modèle naïf	NA	-0.129	-2.51	0	0.0421	-0.397
3	R2_OOS modèle ARIMA	0.115	NA	-2.11	0.115	0.152	-0.237

Annexe 20 : Résultats de la prévision naïve.

```
> show(prev_naive)
```

	Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jan 2019	8.461826	8.402811	8.520842	8.371570	8.552083	
Feb 2019	8.461826	8.378366	8.545287	8.334184	8.589469	
Mar 2019	8.461826	8.359608	8.564044	8.305497	8.618155	
Apr 2019	8.461826	8.343795	8.579858	8.281313	8.642340	
May 2019	8.461826	8.329863	8.593789	8.260006	8.663646	
Jun 2019	8.461826	8.317268	8.606385	8.240744	8.682909	
Jul 2019	8.461826	8.305686	8.617967	8.223030	8.700623	
Aug 2019	8.461826	8.294905	8.628748	8.206542	8.717111	
Sep 2019	8.461826	8.284779	8.638873	8.191056	8.732596	
Oct 2019	8.461826	8.275202	8.648450	8.176410	8.747243	
Nov 2019	8.461826	8.266094	8.657559	8.162479	8.761174	
Dec 2019	8.461826	8.257390	8.666263	8.149168	8.774485	

Annexe 21 : Pr vision na ve.



Annexe 22 : Erreur quadratique moyenne de pr vision mensuelle (MSE en log) par mod le – Ann e 2019

Mois	Na�ve	ARIMA	Holt-Winters	ADAM.ETS	ADAM.ETS.ARIMA	SSARIMA
Janvier	0.00291	0.00478	0.00537	0.00291	0.00425	0.00506
F�vrier	0.0134	0.0170	0.0253	0.0134	0.0137	0.0183
Mars	0.0285	0.0352	0.0581	0.0285	0.0298	0.0406
Avril	0.0562	0.0662	0.1180	0.0562	0.0532	0.0826
Mai	0.0654	0.0776	0.1550	0.0654	0.0570	0.1000
Juin	0.0930	0.1060	0.2300	0.0930	0.0832	0.1340
Juillet	0.114	0.133	0.314	0.114	0.0953	0.170
Ao�t	0.140	0.158	0.405	0.136	0.119	0.201
Septembre	0.169	0.196	0.532	0.169	0.157	0.249
Octobre	0.206	0.236	0.681	0.206	0.196	0.300
Novembre	0.255	0.290	0.870	0.255	0.241	0.365
D�cembre	0.310	0.350	1.090	0.310	0.297	0.433

Annexe 23 : Résultats du test de Diebold-Mariano.

```
df_dm_pvalue
A tibble: 1 × 5
  ARIMA Holt_Winters ADAM_ETS ADAM_ETS_ARIMA SSARIMA
  <dbl> <dbl> <lg1> <dbl> <dbl>
1 0.000000133 0.000643 NA 0.828 0.00000396
```

Annexe 24 : Prédiction à un pas.

